UNDERPINNING AND ADVANCING COGNITIVE ENGINEERING OF INTELLECTUALIZED CYBER-PHYSICAL SYSTEMS

Dissertation for the doctoral degree of the
Hungarian Academy of Sciences

Imre Horváth

Budapest, 2023

This dissertation book and the accompanying booklet of propositions are based exclusively on genuine academic research results and pure human semantics, and have been digitally processed without using any supplementing software means, such as generative artificial intelligence tools, or chatbots generated synthetics text, images, and/or algorithms or other digital additives.

Table of contents

Chapter 1

1. Prologue

1.1. Setting the stage

We are living in an age which is characterized by a dominant extreme acceleration, in particular, in the field of technologies and engineering (Gosling, 2020). One indicator of the acceleration is shortening of the time periods of the emergence, culmination, and dissolution of subsequent industrial and social revolutions (Figure 1.1). The first publications about the fourth industrial revolution appeared in the mid-1980s, and reported on computerization and informatization of productive, managerial, and administrative industrial processes (Schwab, 2017). Though this trend continues to proliferate, many authors have reported on the advent of the fifth industrial revolution, which is being triggered by the revitalization of **artificial intelligence research**, the results of which permeate not only the industry, but the whole society, and trigger the propagation of tools and systems exploiting various forms of artificial intelligence (AI) (Pathak et al., 2019). While the first three industrial revolutions aimed at extending the physical capabilities of human beings, the fourth and fifth revolutions are orientated to augmenting informative and cognitive potentials of humans through systems engineering. The industrial revolutions brought about not only technological sophistication, but also growing complexity and heterogeneity of engineered systems.

The intention of making intelligent systems has emerged well before our digital age. Nonetheless, **intellectualization of engineered systems** has become a probability and a possibility only at the time of the emergence of digital computing (Baier et al., 2023). Depending on the level of reproducing/mimicking the operation of the human brain, system intellect can be classified according to five levels, known as reflexive, imperative, adaptive, autonomous, and cognitive intellect. This is in line with the well-known levelling of human

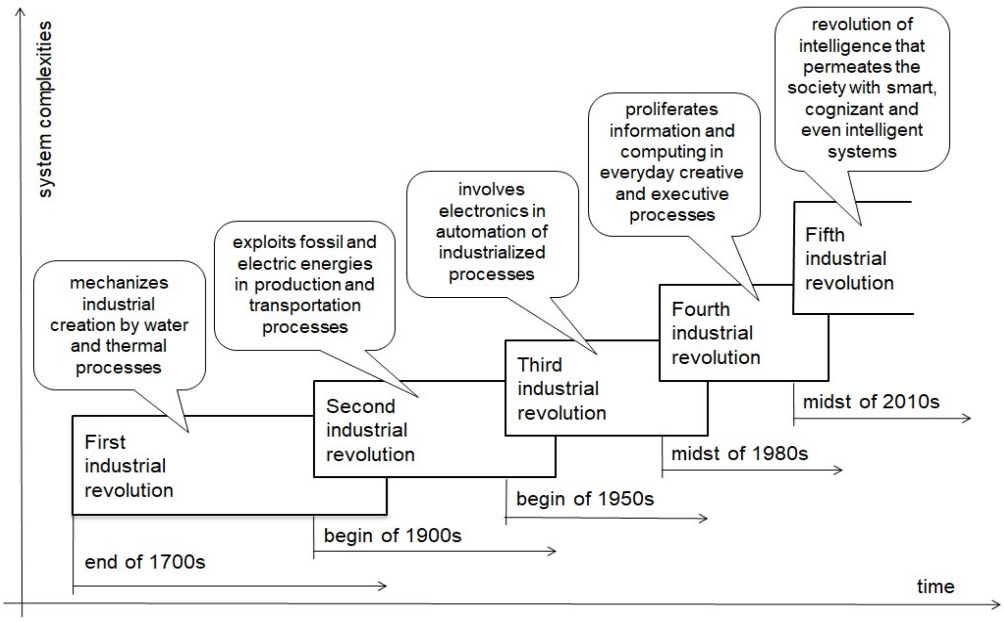

Figure 1.1: Orientation and shortening periods of industrial revolutions (in a quasi-logarithmic scale of time)

intelligence. In this work, intellectualization of engineered systems is tackled as the process of equipping them with capabilities for computational problem-solving, i.e., making them able to possess, explore, aggregate, synthesize, share, and exploit knowledge and to reason with it towards fulfilling operational objectives (Miki and Yamakawa, 2008). In this sense, intellect is a computational subset of the overall human intelligence. However, intellectualization is a broader concept than rationalization. While rationalization works with logical constructs and symbolic mechanisms in computation, intellectualization works with semantic constructs and cognitive mechanisms. Rationalization was typical for the early knowledge-based systems (Baldassarre and Granato, 2020). Intellectualization is conjointly about the rational content (system knowledge) and the transformational agents (reasoning mechanisms). Intellectualized systems show various self-* features, such as self-awareness, self-learning, self-reasoning, self-control, self-adaptation, self-evolution, and self-reproduction. These features make them nondeterministic, goal-orientated, context-aware, causality-driven, reasoning-enabled, and retrospectively, inductively or abductively learning systems.

Notwithstanding these, making truly intelligent systems has remained a tremendous challenge as well as an open issue for our days (Sheppard, 2019). Over the last six decades, a lot of efforts have been invested into artificial intelligence and various approaches have been proposed for its conceptualization and implementation. As a result of the world-wide intense efforts, the '*terra incognita*' of artificial intelligence is gradually converted into a '*terra quaestuosa'*, and even beyond. (Here, by using the Latin phrase '*terra incognita*', I denote an unexplored field of knowledge, while by using the term '*terra quaestuosa*' - literally means 'problem land' – I refer to the stage of progress that do need further inquiries and realizations in order to know more about what is probable and what is possible). However, cognition, intelligence, and systems research and engineering have not reached any culmination points yet, though some writers claim that they have already reached the so-called '*terra utilis*', where the developed tools and systems have irreversible effects on the fundamentals and operation of science, economy, and society. In spite of the philosophies and forecasts mushrooming in current literature, the fact of the matter is that there is more ahead of us than behind in this field in all respects (Commuri et al., 2018).

The field of AI is not unified by a shared theoretical foundation or a common goal, but by a class of loosely related problems. Research and development in artificial intelligence have spread over more than 100 domains of interest, without any shared theoretical foundation, common goals and criteria, or integrative conceptual and methodological frameworks. Wang (2006) stated that, as enabled by the current computational resources, artificial intelligence is not what it should be. He argued that the mainstream works in the field are on domain-specific perception and cognition problems and solutions, whereas AI is supposed to focus on general-purpose systems that are adaptive to their environments, and can work with insufficient knowledge and resources. I share his opinion that a complete AI research should result in: (i) a theory on the principles and mechanisms of intelligence, (ii) a formal model of intelligence based on the above theory, and (iii) a computer implementation of the above model.

1.2. Scientific interests and objectives

Introduced some fifteen years ago in the USA, the **paradigm of cyber-physical systems** (CPSs) is being implemented in many different forms and for many practical applications world-wide (Gill, 2006). After the initial euphoric excitement, ungrounded hype, and failed predictions concerning their key role in Industry 4.0, research and development of CPSs has come to a normal period, which is driven by rational and critical transdisciplinary thinking. As evidenced by the recent literature, not only current possibilities, but also near-future ones are widely studied. In the meantime, many of the objectives formulated as progressive fifteen years ago, such as "use computations and communication deeply embedded in and interacting with physical processes to add new capabilities to physical systems", has become quite insignificant by now. The paradigm is paving its own way further based on much broader theoretical and technological fundamentals. I strongly believe that our near future will be influenced deeply by the paradigm of CPSs, in particular by intellectualized, socialized, and personalized CPSs.

Erwin Schrödinger wrote: "The task is not much to see what no one has yet seen, but to think what nobody has yet thought about that which everybody sees." Having this inspiration, the primary goal of this dissertation is to present the outcomes of multiple interrelated studies and to synthesize fundamentals, frameworks, models, and methods worthwhile for both follow-up inquiries and practical developments. Towards this end, it reports on the investigations and propositions concerning the conceptual underpinning and methodological advancement of **cognitive engineering of CPSs**. More specifically, the dissertation addresses and elaborates on three interrelated domains of overall interest, namely (i) evolution of CPSs as a consequence of their growing intellectualization, (ii) studying and managing synthetic intellect (self-generated system knowledge and ampliative reasoning mechanisms) in a systematic manner, and (iii) dealing with operational opportunities and issues opened up by the affordances of intellectualization.

The research presented in this dissertation has made an attempt to learn and construct a better understanding of a number of frontier issues that the discipline of CPSs design and engineering has been facing in the last decade and will presumably be facing in the coming decade. As the title indicates, the completed research has focused on the underpinning and advancing cognitive engineering of intellectualized CPSs and has tried to contribute to systems science and engineering. More specifically, it has dealt with **synthetic system-level knowledge** that is the key constituent of the synthetic system intellect of CPSs, together with the associated ampliative computational mechanisms. Understanding its essence and facilitating its utilization within and across CPSs were in the focus of the work, rather than development and application of specific systems. Therefore, the work and its significance need to be assessed from the perspective of road-mapping in systems science, systems thinking, and systems engineering.

Some of the research activities concentrated on phenomena and problematics (i.e., uncertainties and difficulties inherent in challenging problems and complicated situations) that still belong to the *'terra incognita'*, while others addressed those that are parts of the *'terra quaestuosa'*. The results are theoretical and methodological in nature, though they have been used in the development of certain **demonstrative prototype systems**. For the sake of clarity, having a different orientation and purpose, the work presented in this dissertation is not per se a contribution to the research and development of artificial intelligence. Nevertheless, some parts may be regarded as input for specific narrow fields of artificial intelligence research and utilization - as much as artificial intelligence research, development, and implementation have been playing a triggering and enabling role in intellectualization of engineered systems.

At the focus of the dissertation are **intellectualized cyber-physical systems** (i*CPSs), which are often and interchangeably also referred to as smart cyber-physical systems (s*CPSs) (though the latter is just a subset of the former). Their current representatives are such as self-driving vehicle systems, medical operation theater systems, home care robotic assistance systems, and natural disaster avoidance systems. The dissertation discusses a number of recognized concern-domains that are relevant and important from the perspective of advancing to next generations of cyber-physical systems (NG-CPSs), which are supposed to be equipped with substantial intellectual capabilities. From the many potential concern-domains, the research underpinning this dissertation concentrated on seven important phenomena and problematics, respectively.

1.3. Overview of the contents

The contents of the dissertation interlink and blend the outcomes of many related research works of contemporary literature and the own results of the author, which have been published previously. These concern both the fundamentals of CPSs and a generalizable interpretation of what system intellect is with a view to natural human intelligence and artificial narrow intelligence. The focus is on i*CPSs for everyday operation and servicing. The novelty of the results can be judged easily by comparing them with the contemporary literature. Many of them open up new trajectories of investigations and raise the need for follow-up research to extend and/or consolidate the explanations and specifications. In this context, usefulness has been preferred to completeness.

Structurally, the dissertation is organized into nine chapters. It includes an introductory and a conclusive chapter, called Prologue and Epilogue, respectively. The main body is formed by seven chapters that present and discuss the findings and results of both the explorative and constructive research activities. The titles and the order of the chapters are shown in Figure 1.2. The first chapter elaborates on the different types of definitions of CPSs and on the **paradigmatic features** of non-intellectualized and intellectualized CPSs, respectively. The second chapter (i) presents a **model of progression** of CPSs, (ii) introduces the concept of system generations, and (iii) characterizes the various generations of CPSs. The third chapter explains the unique nature of **system-level synthetic problem-solving knowledge** and demarcates it from the knowledge cultivated by alpha, beta, gamma, and delta sciences. Considering the current progress of generating synthetic system knowledge by means of, for instance, specific computational learning mechanisms, such a study is not only timely, but also indispensable from the point of view of further utilization of synthetic system knowledge.

Reasoning with the different objectives of gnoseology and epistemology, the fourth chapter introduces **sympérasmology** as a possible conceptual framework and investigation approach of system-level problem solving knowledge. The fifth chapter clarifies the need for **prognostic systems thinking** (PST), and discusses its conceptual pillars and analysis concerns from the perspective of i*CPSs. The sixth chapter elaborates on the **aggregation, transfer, and exploitation of synthetic knowledge** self-acquired and/or self-generated by intellectualized systems as an industrial asset. Starting out from the co-emerging trends of complexification,

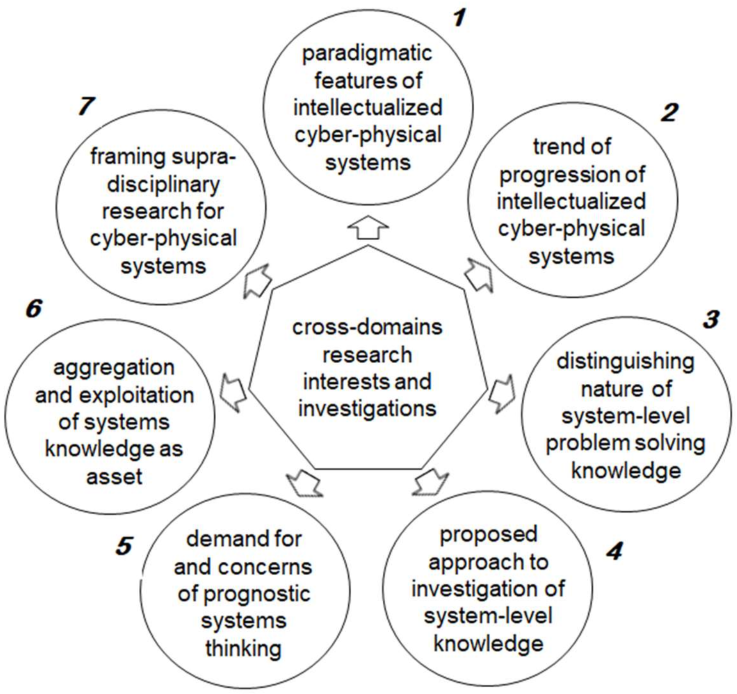

Figure 1.2: The research topics addressed in the chapters of the dissertation

intellectualization, socialization, and personalization of CPSs, the seventh chapter addresses the nature of **pluridisciplinary research** approaches and proposes supradisciplinary research as a proper and necessary organizational and methodological framework for doing research in complex problematics and phenomena related to next-generation CPSs. The contents of the chapters are interrelated by the cross-domain research interest and investigations. The major scientific contribution of the discussed work concerns the novel disciplinary domain of cognitive engineering of systems. Parts of the remaining sections of this chapter have been compiled by reusing contents from the following peer-reviewed publications: H1, H2, H3, and H4 (see Appendix A.1.1).

1.4. Cognitive engineering of systems

Owing to the results of system science, artificial intelligence, cognitive engineering, and complex adaptive systems research, the family of intellectualized systems is becoming more and more powered by context-dependent knowledge and reasoning mechanisms (Lungarella et al., 2007). From the viewpoint of the knowledge possessed by i*CPSs, not only its human-embedded part plays an important role, but also the part that is self-acquired or self-synthesized by these systems. The additional chunks or bodies of human knowledge (and the related computational mechanisms) are supplemented incrementally to the possessed knowledge and mechanisms of an i*CPS at the discrete times of system updates or upgrades. The self-acquired knowledge and computational enablers of the system may evolve and become operationalized continually. Eventually, system knowledge that complements common-sense and scientific human knowledge is maturing into a **productive asset** for the various industries and for segments of the society (Woods, 1987).

Cognitive engineering of systems (CES) is a relatively novel domain of scientific and professional interests (Wilson et al., 2013). It has been brought about by the need for systematic investigation and implementation of the (i) cognitive enablers, (ii) technological realizations, (iii) social relationships, and (iii) application opportunities of current and next generation systems. CES supports (i) equipping i*CPSs with application-specific knowledge, (ii) constructing computational reasoning mechanisms that elaborate on that knowledge, and (iii) embedding i*CPSs in real-life application environments. These together enable i*CPSs to (i) build awareness of the dynamic context of their operation and application, (ii) infer not pre-programmed 'intellect' (data, information, knowledge, meta-knowledge) based on monitoring the states of the problems, their performance, and the embedding environment, (iii) reason about the goals and change them, (iv) develop and validate plans for reorganization to meet the goals of working, (iv) adapt, evolve, and reproduce themselves on their own as needed by a better operational performance, (v) enable the interaction with and among i*CPSs on multiple levels, and (vi) share tasks with and among i*CPSs and people in various contexts. In addition to designing genuine core functions of operation and behavior, such as awareness building, abductive reasoning, operation forecasting, apobetic interaction, and self-managed adaptation, and optimizing 'human-in-the-loop' and 'system-in-the-loop' situations, CES is also interested in the interlinked processes of system intellectualization and designing physical, cognitive, and affective interactions of humans with intellectualized systems, and vice versa.

Notwithstanding, the 'science' of **self-managing systems** is still in a premature stage of development and is deemed to be a rather unsettled domain of inquiry. In particular, this applies to self-management of system knowledge (Couch, 2023). An important issue is that neither an overall theory of system-level problem solving knowledge of intellectualized systems, nor a comprehensive methodology for purposeful intellectualization exists. This makes cognitive engineering of systems tentative and experimental. Some researchers suggest that time has come to establish a philosophically underpinned theoretical framework. This motion is seconded by the on-going intelligence revolution, in which artificial intelligence becomes a productive power, an enabler of smart systems, and a strong transformer of social life. One of the novel contributions of this dissertation is a skeleton of the needed generic theory of system knowledge (and a possible new branch of philosophical studies). In full awareness of the accelerated scientific changes and the fast pace of technological innovations, it also deals with the issues of prognostic systems thinking and the conversion of distributed systems knowledge into a common asset.

1.5. On the overall research approach

A holistic theory but sufficiently articulated theory of CPSs is still in an early stage of development. Though the disciplinary field is a multidisciplinary or even transdisciplinary, the included disciplines (e.g. mechanics, electronics, computing, networking, and informatics) keep their knowing and working traditions and cultures. Instead, integrative and collaborative **supradisciplinary research** strategies are needed that purposefully interconnect monodisciplinary, interdisciplinary, multidisciplinary, and transdisciplinary approaches, methods, and knowledge (Funtowicz and Ravetz, 1993). In this dissertation, the idea of transdisciplinary research integration has been interpreted in line with the basis of prior literature, along with similar terms such as interdisciplinary and multidisciplinary research (Giri, 2002). Attempting a seamless procedural synthesis of these, supradisciplinary research would open up a multidimensional space of inquiry that is characterized by (i) concurrent dependence on multiple (physical, biological, human, social, computational, technological, etc.) domains of inquiry and investigations, (ii) various progression levels (discovery, description, explanation, prediction, and regulation) of knowledge with regard to the studied phenomenon (and problematics), and (iii) the need for synergy in terms of hardware, software, cyberware, brainware, etc. related knowledge. Evidently, such a methodological approach cannot be realized by one single researcher or a relatively small team of researchers. On the other hand, the completed research work features a broad perspective and a transdisciplinary flavor in the sense that the generated knowledge is above the level of the individual domains and can be projected to many of them (Horváth and Pourtalebi, 2015).

At the outset of the research work, **two fundamental principles** were considered: (i) any attempt to specify what intellectualized systems are must start with defining what intelligence (and in particular system intelligence) is, before dealing with the process and forms of intellectualization, and (ii) the basis of a intellectualization of engineered systems is, first of all, human knowledge, but the growing amount and role of problem solving knowledge synthesized by system(s) must be taken into account at considering their near-future manifestations. It is well-known for scientists that human intelligence is an extremely complex phenomenon with multiple appearances and we have arrived at neither a canonical interpretation nor a consensus definition yet. The diverse range of interpretations of the same concept causes more confusion than clarification.

To have a starting point and to underpin the investigations, let us take the definition proposed by Gottfredson (1997) which says: (human) “intelligence is a very general mental capability that, among other things, involves the ability to reason, plan, solve problems, think abstractly, comprehend complex ideas, learn quickly and learn from experience. It is not merely book learning, a narrow academic skill, or test-taking smarts. Rather, it reflects a broader and deeper capability for comprehending our surroundings – ‘catching on’, ‘making sense’ of things, or ‘figuring out’ what to do was preferred. In full agreement with Monett and Lewis (2020a – 2020b), my concern has been how we can construct intelligence-based systems and comprehend their behavior if we do not or cannot accurately define (i) what cognitive capabilities they have to provide, (ii) how intelligent behavior may be realized, and (iii) how their intelligent performance can be assessed (measured) and enhanced. On the other hand, if the associated concepts are well defined, there is an opportunity for better understanding, robust investigations, and dependable implementations.

In terms of its nature, the conducted research is combined secondary exploratory research (identifying new phenomena or problematics for research) and tertiary constructive research (conceptualization of theories for the studied phenomena and resolutions for problematics). An example of the former is the theory of CPS evolution through generations, while an example of the latter is a set of concerns about the operationalization of PST. These are discussed in detail in the related later chapters of the dissertation. Various methods were used in the completed research, such as (i) multi-focal literature review, (ii) critical systems thinking, (iii) computational modeling and simulation, and (iv) implementation of testable prototypes of modules and subsystems. The sources of input information were (i) openly accessible journal papers, (ii) professional books, (iii) conference proceedings and presentations, (iv) personal communications with international experts, (v) contents of Internet repositories and forum places (in particular, Research Gate and Academia), and (vi) collaborative work with staff colleagues, and Ph.D. researchers.

The outcomes of different studies were also the subject of philosophical and methodological speculations. Due to these multi-methodological approaches and treatment, the overall character of the work and the dissertation is explorative, investigative and argumentative, rather than prescriptive or experiential. The flow of reasoning and argumentation meanders through theoretical concepts and practical facts, but also includes narrative thinking and untested personal opinions. It is hoped that the addressed topics, including the white and grey spots of knowledge, indeed deserve attention, or even a broader public debate. It is also expected that, towards a collective consolidation, follow-up works and discussions will extend the contents, resolve the open issues, provide more empirical and rational evidence, and rectify the factual slips and misunderstandings.

Chapter 2

2. Paradigmatic features of intellectualized cyber-physical systems

2.1. Research objectives and approach

The main objective of this chapter is to provide a concrete distinguishing characterization of i*CPSs without introducing mathematical abstractions and symbolic constructs (Willems, 1991). The need for this was recognized from the perspective of public communications. As stated in the first chapter, the realm of engineered systems is wide and, evidently, all families of the implemented systems are different. This also applies to intellectualized systems. Figure 2.1 displays the most significant characteristics of engineered systems. As the double-arrowed lines indicate, the major characteristics usually appear as opposites of each other. By an apt combination of these characteristics, a system can be described with such a fidelity that uniquely distinguishes it from even resembling other systems. However, this is not as simple as it seems at first sight. There are three reasons for it: (i) the paradigm of CPSs has grown out from multiple well-established disciplines, such as embedded systems, mechatronics and robotics systems, complex adaptive systems, collaborative multi-agent systems, Internet of things systems, and hybrid information processing systems, and many of the definitions of CPSs reflect these roots, (ii) in addition to functional/structural complexity and variety, also a disciplinary heterogeneity is present in these systems, which has actually been further increased by the appearance and involvement of disciplines such as artificial intelligence, big data science/engineering, cognitive science, and social systems science, and (iii) the paradigm is being made dynamic by the concurrent intense increase and convergence of scientific knowledge, acceleration of technology synthesis, and widening of the application purposes and domains.

The content of this chapter has been compiled from the following peer-reviewed

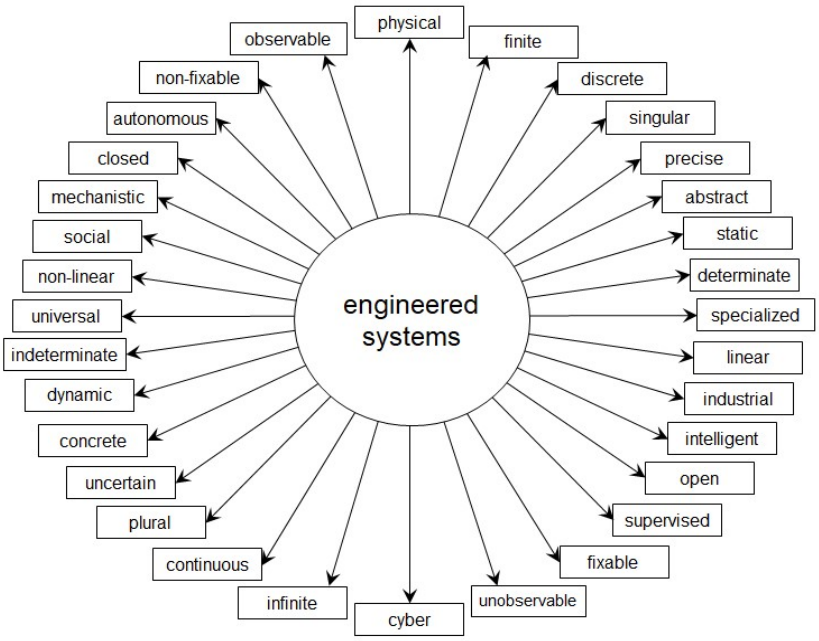

Figure 2.1: Major characteristics of engineered systems

publications: H5, H6, H7, and H8 (see Appendix A.1.1). Considering the disciplinary convergence and the progress in technology integration, the interconnected goals of this investigation are (i) to get deeper insight into the essence and foundations of CPSs, (ii) to identify their paradigmatic features that distinguish them from other engineered systems, (iii) to cast light on the typical functionalities of CPSs, (iv) to make an inventory of the enabling technologies, and (v) to analyze their intellectualization from the point of view their behavioral characteristics. This study was stimulated by the fact that a large number of differing definitions which can be found in the literature are based on different interpretations and conceptualizations. Paradigmatic (notional modeling) features of CPSs have been deemed to be crucial elements of their identification, classification, and characterization.

2.2. Cyber-physical systems from a birds-eye view

The paradigm of CPSs emerged less than two decades ago. In the context of the 4th industrial revolution, they have been regarded as major technological and infrastructural enablers for productive and transformative industries (Ahmad et al., 2016), as well as crucial resources for the implementation of smart interconnected society (Colombo et al., 2017), contributing by informatization, integration, and automation. In the context of the 5th industrial revolution, they are seen as the basis of intellectualization, socialization, and humanization of engineered systems for societal and personal purposes. Technologically, this system paradigm implies a distinct category of complex, heterogeneous, data- and software-integrated, and intellectualized engineered systems. One trait of these engineered systems is that they complementarily operate in the physical, cyber, and mind realms, and perform various forms of self-control, self-adaptation, and self-evolution. It is a fact that the cyber world can gradually penetrate not only into the unanimated and animated natural worlds, but also the human perception, cognition, and affective domains through various miniaturized synergistic and organic technologies. By extrapolating from the latest progression in the field of complex systems, it can be claimed that the near-future (next-generation) CPSs will largely be different because of their growing self-intellectualization and self-evolving capabilities (Weyns et al., 2021).

Current industrial CPSs integrate signals and data obtained from dynamic and uncertain environments with data-driven software control, context-sensitive decision-making, and continuous physical changes (Gerritsen and Horváth, 2012). It has been recognized that, both in the society and in the industry, there will be a growing demand as well as numerous opportunities to utilize interacting CPSs, which are equipped with strong connectivity and collective problem solving abilities (Ansari et al., 2018). The CPSoS Consortium specified CPSs as complicated systems that exhibit the features of systems of systems, i.e., (i) large, often spatially distributed physical configuration with complex dynamics, (ii) distributed control, supervision, and management, (iii) partial autonomy of the sub-systems, (iv) dynamic reconfiguration of the entire systems on different time-scales, (v) continuous evolution of the systems during their operation, and (vi) possibility of emerging behaviors (CPSoS Consortium, 2015). In the view of (CPSoS Consortium, 2016), the majority of NG-CPSs will manifest as cyber-physical systems of systems (CPSoSs).

It must be mentioned that the different conceptualizations and different vocabularies of the involved disciplines create (i) particularity in specifications, (ii) ambiguities in communication, and (iii) borderlines in collaboration. Many publications have interpreted CPSs as: (i) functionally and technologically extended embedded systems, (ii) sophisticated implementations of Internet of things systems, (iii) augmented advanced mechatronics systems, or (iv) complex collaborative adaptive systems. For instance, Quadri et al., (2015) suggested that CPSs are next-generation embedded systems and explained this with the accelerated development of sensing, networking, and communication technologies. Considering the number of shared functions, other researchers have argued that CPSs are not radically different to the Internet of things systems (Xu et al., 2018), or to the proactive digital twins (Koulamas and Kalogeras, 2018).

Surprisingly, certain researchers emphasize loose connections - rather than tight functional and architectural connections or synergy - among the system's constituents. According to Tan et al, (2008), "cyber-physical systems are a next-generation network-connected collection of loosely coupled distributed cyber systems and physical systems monitored/controlled by user-defined semantic laws" and whose prototype architecture supports design goals such as (i) global reference time, (ii) event/information driven communication, (iii) quantified confidence indicators, (iv) publish/subscribe scheme, (v) semantic control laws, and (vi) secure networking technologies. Liu et al. (2017) argued that: "In fact, all defense systems (such as aircraft, spacecraft, naval vessels, ground vehicles, etc.) and subsystems in those systems are all CPS. Additionally, integrated circuits, micro-electro-mechanical systems (MEMS) and nano-electro-mechanical systems (NEMS) also belong to CPS." These issues will be resolved in the rest of the dissertation.

2.3. Cyber-physical systems as results of convergence

Our age is characterized by the concurrency of convergence and divergence, and their dynamic interaction. It means that while certain disciplines are combining their inquiry objectives, methods, and knowledge, completely new investigation domains emerge within the integrated disciplines. By merging mechanics, electronics, and computing, mechatronics has been one of the typical examples of disciplinary convergence. Within this transdisciplinary branch of knowing and making, the coexistent divergence is exemplified by the abundance of research topics and approaches that belong to or form a part of none of the mentioned original disciplines. The National Research Council of the USA interpreted convergence as forms of domain-diverse research to mean an "approach to problem solving that integrates expertise from life sciences with physical, mathematical, and computational sciences, medicine, and engineering to form comprehensive synthetic frameworks that merge areas of knowledge from multiple fields to address specific challenges" (NRC, 2014). Actually, this has happened as a combinatorial innovation in the field of CPSs. Notwithstanding, development of these systems is a transdisciplinarity challenge, implying the need for synthetic paradigms, comprehensive frameworks, and interests across boundaries. (Gooding et al., 2022).

At the beginning of the 1970s, it was recognized that innovative and competitive systems, artifacts, and processes could not be developed relying on just one or two single disciplines. This has given impetus to the investigation of the foundations of convergence (by cross-disciplinary operations) and to facilitate technology-intensive and -sensitive constructive disciplines (by systematic meta-fusion of knowledge). Lipscomb et al. (2023) has characterized the 21st century as the era of rising convergence engineering, which operates on three levels: (i) at macro-level (where it concerns integration of philosophical platforms, grand theories, paradigms, and models), (ii) at meso-level (where it addresses organization of domain-diverse research projects that focusses on cultural homogenization, knowledge integration, and methodological cross-fertilization), and (iii) at micro-level (where it deals with the individual and team attitudes and competencies towards cross-disciplinary collaboration and co-creation (Wetter, 2006). In this interpretation, convergence is a problem-solving approach on strategic, tactical, and operational levels. The meso-level interconnects translational research (transferring scientific knowledge into new systems technologies) and transformative research (transferring technological generalizations into new scientific models) (Xue et al., 2016).

This dissertation sees CPSs as one of the early tangible results of scientific and technological convergence. Specifically, the convergence of information technology, nanotechnology, biotechnology, cognitive sciences, and social sciences with conventional systems science, artifact engineering, and material/process technologies is paving the way to disruptive implementations. The integration culminates in fusing bits, atoms, neurons, genes, and memes (Figure 2.2). This is often referred to as the **bits-atoms-neurons-genes-memes** revolution (BANGM) (Horváth and Tavčar, 2021). It is a fact that the existing CPSs represent practical examples of the integration of bits and atoms in human and social contexts. The phenomenon of integration of atoms, bits, and neurons is exemplified by the emergence of cyber-biophysical systems (represented by assistive and corrective implants and artificial limbs/augmentations) and by the showcased results of gentelligent systems, biological analogies-based, and physically-driven neural network systems (Kumar et al., 2021).

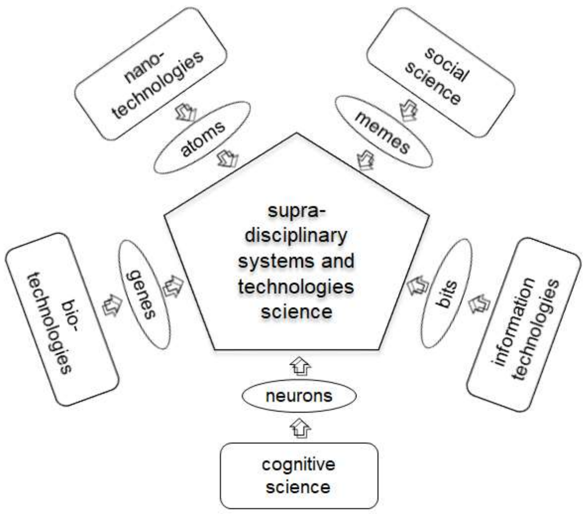

Figure 2.2: Essence of the BANGM revolution

This dissertation anticipates further disruptive achievements and drastic effects of the continuing technology synthesis on the possibilities of developing ‘disappearing’ (invisible) CPSs (Rai and Rai, 2015). Furthermore, it is believed that the growing interest in integration of neurons, genes, and memes will support extensive socialization, personalization, and personification of various industrial and non-industrial systems. As elements of a social culture and personal behavior passed from one individual to another by non-genetic means (e.g. learning, imitating, etc.), memes are digitalized by images, videos, sounds, notes, signs, rules, patterns, emoji’s, and other constructs.

2.4. Epitomizing definitions of cyber-physical systems

The science of CPSs is still in the formation stage and the procedural approaches and methodologies are heavily influenced by the reductionist and disciplinary thinking of the past. The publications on the theoretical fundamentals typically elaborate on domain theories concerning: (i) system control, (ii) system modelling, (iii) system architecting, (iv) system communication, (v) system security, and so forth. Only a few efforts have been reported in the contemporary literature concerning the development of comprehensive multidisciplinary (or transdisciplinary) theories for this family of systems. These theories can reduce the notional and methodological uncertainties (Chen, 2017). Nevertheless, practitioners have questioned their usefulness from the viewpoint of supporting their practices.

The first step towards a general theory of CPSs is a comprehensive but discerning ontological definition that states what exists in the form of these systems. However, for the above-mentioned reasons, finding a robust definition of CPSs is not as obvious as it seems at first sight. One of the early definitions, introduced by Lee (2007), claimed: “cyber-physical systems are integrations of computation and physical processes”. Over the years, several more articulated definitions have been proposed that reflect the newer scientific concepts, technologies, and functionalities. As though the current literature has not parsed these definitions in detail, this has been a primary objective of our exploratory research. It has been found that the variety of the published formal definitions (specifications) could be classified into six categories, namely: (i) augmentative, (ii) descriptive, (iii) normative, (iv) predictive, (v) symbolic, and (vi) domain specific definitions.

The **augmentative definitions** express in which sense CPSs are more than other comparable (more traditional) systems. That is, starting out from the characteristics of systems such as embedded systems, real-time systems, network-based systems, etc., they identify the characteristics which distinguish CPSs from the other system categories mentioned above. An example is the definition formulated by Marwedel (2021), which posited that CPSs are embedded systems plus a (dynamic) physical environment. The appropriateness of this definition is explained by the fact that sensors, actuators, and processors are often embedded in the physical parts of these systems (e.g., gentelligent systems) as well as in their direct environment.

The **descriptive definitions** try to bring all (or at least a large number of so far experienced) common characteristics of CPSs into an exact holistic formulation in a specific context. A typical descriptive definition is: “In a cyber-physical system, a physical mechanism is controlled or monitored by computer-based algorithms, the physical and software components are deeply intertwined, able to operate on different spatial and temporal scales, exhibit multiple and distinct behavioral modalities, and interact with each other in ways that change with context” (Putnik et al., 2019).

The **normative definitions** intend to capture a minimal set of criteria that specific implementations of CPSs should meet in order to be regarded as such. A typical normative definition claims that CPSs are supposed to: (i) manifest as complicated networked multi-actor systems, (ii) implement multiple sensing-reasoning-learning-adapting loops, (iii) be realized normally as synergistic system of systems, (iv) be tailored to service provisioning and dynamic resource management, and are characterized by (v) deep penetration into real-life physical processes, (vi) use data and patterns driven cyber-physical computing, (vii) capability to exploit a growing level of system intelligence, and (viii) provide benefits in applications in human, social and industrial contexts (Horváth, 2014). The specification of the CPSoS Consortium cited in the previous section is also a normative definition.

The **predictive definitions** include tendencies, projections or perceptions to forecast distinguishing characteristics of future (next-generation) CPSs. For instance, (Dumitrache, 2010) foresees that “The next generation of CPSs will integrate hardware and software designed by integration of dynamical properties of physical objects and advanced control strategies or even intelligent hybrid methodologies. The change in conception of CPSs will create new architectures of large-scale intelligent CPSs with new capabilities and improve the quality of production and life.” On the other hand, (Robinson, 2021) believes that “Future CPS will function like teams, where subsumption occurs with CPS merging and separating. They become one CPS, not because everything is managed centrally, but because they are collaborating as a team, sharing some higher levels of supervision. For instance, in terms of the supervision of safety, if an accident occurs, the law searches for one entity to take responsibility.” Though predictive definitions carry uncertainties, they are useful means to support vision forming or road mapping.

The **symbolic definitions,** used in systems science and systems related theories, create canonical models or constructs in order to capture the essence of a family or an instance of a CPS using logical, mathematical, and information technological means (Togay, 2014). Due to the complexity and heterogeneity of CPSs, symbolic definitions and mathematical models are typically partial. For instance, interpreting an anthropomorphous robot as a CPS, Stepanov et al. (2020) proposed a set-theoretic model of the hardware-software complex of the architecture of a generic CPS in the following way:

$S = \langle F, A, G, Rl, T, D, O \rangle$

where $F = \{f_i\}$ is a set of functional components, $A = \{a_i\}$ is a set of algorithms of component functioning $\dim(A) = n_a$; $G = \{ \langle f_i, f_j, r_{ij}(f_i, f_j)\rangle | f_i \in F, f_j \in F, r_{ij}(f_i, f_j) = \{ 0|r_k \}, r_k \in Rl, n_R = \dim(Rl), k = 1, n_{R,i} = 1, n_{f,j} = 1, n_f \}$ is a scheme (graph) of component interrelations; $Rl = \{r_1, r_2,..., r_{nR} \}$ is a set of component interrelations for the system; $D = \{ d|d = \langle p, c, m\rangle \}, d \cdot m \in C^{nm1} \times C^{nm2} \times C^{nm3} \times C^{nm4}, d \cdot c \in C^{nc1} \times C^{nc2} \times C^{nc3}, d \cdot p \in$ {true, false} is a set of types of data information structures used for solving the tasks, where C^x is a x-dimensional space of complex numbers; $P = \{ p_i|p_i \in \{\text{true, false}\}, p_i = f_{pi}(d_x, d_y)\ d_x, d_y \in D \}$ is a set of the relations used in settings of the tasks being solved; $O = \{o|o = \langle o_c, o_s, o_r, o_d\rangle, o_s = \{ d|d \in D\}, o_r = \{d|d \in D\}, o_c \subset \{\{\{d \cdot p\}^*\} \times Rl\}, o_d \subset Rl \}$ is a set of actions (operations) used for solving the tasks; $T = \{ t|t = \langle t_S, t_R, t_D\rangle\}, t_S \subset D^*, t_R \subset D^*, t_D \subset \{\{\{d \cdot p\}^*\} \cup Rl \}, d \in D\}$ is a set of tasks, where $X^* = X^0 \cup X^1 \cup X^2 \cup ... \cup X^N \cup ..., X^0 = \emptyset$, $t_S = S_{rc}(t)$ are input data, $t_R = R_{qr}(t)$ are required results, $t_D = D_{mnd}(t)$ are requirements to the results of a $t \in T$ task solution.

The **domain-specific definitions** specify CPSs that are concurrently influenced by the generic paradigm and the specificities of the target application domain. These domains are such as (i) manufacturing, (ii) transportation, (iii) agriculture, (iv) horticulture, (v) delivery, (vi) urban, (vii) medical, (viii) homecare, (ix) sports, (x) crowd management, and so forth. They may be considered both broadly or narrowly. For instance, Deka et al. (2018) derived a specification for transportation CPSs as a generic descriptive definition so as "systems with embedded software (as part of devices, buildings, means of transport, transport routes, production systems, medical processes, logistic processes, coordination processes and management processes), which: (i) directly record physical data using sensors and affect physical processes using actuators, (ii) evaluate and save recorded data, and actively or reactively interact both with the physical and digital world, (iii) are connected with one another and in global networks via digital communication facilities (wireless and/or wired, local and/or global), (iv) use globally available data and services, and (v) have a series of dedicated, multi-modal human-machine interfaces". Complementarily, Sampigethaya and Poovendran (2013) specified an aeronautics transportation CPS so as "it (i) has self-monitoring and self-correcting aircrafts, (ii) autonomously optimizes and supports decision-making in all aspects of fuel efficiency, (iii) inter-flight separation during landing, take-off, or in-air to optimizing operational revenue, and (iv) providing a personalized experience to passengers which will include their desired relaxing/working environment while on board or at airports".

Deemed to be the most representative, the discussed attempts provide a formwork for an evidence-based and formalized definition of the essence of CPSs. They also give a basis for the following conclusions: The published definitions are subjective and reflect different viewpoints and perspectives. Their narratives typically capture generalizations and abstractions (with the exception of the symbolic definitions). They are disconnected from the physical implementation of systems, and are not sufficiently concrete to reflect the diverse purposes of systems. Moreover, they are incomplete and have limitations in terms of rendering the entirety of systems, and do not propose instruments for a systematic comparison of systems. Their actuality is not unconditional, and, in reality, largely depends on the stage of scientific understanding, technology development, and system engineering principles. Equally important is that the majority of the definitions are narrative circumscriptions, which are composed of a set of qualitative linguistic constructs (descriptive characteristics), but usually lack specific quantifiable performance indicators and relationships. This is why additional means of specification have been sought for.

2.5. Paradigmatic system features and profiles

Recently, Putnik et al. (2019) posited that "... a more demanding definition of CPS requires features that both physical and computational system affects and change each other, making a system with systemic relationship, implying a totally new implementation paradigm, and for which the former definition is just a special case. ...". Having recognized this alternative way of characterizing CPSs, our research has targeted a more expressive means of capturing their essence and found it in the form of a paradigmatic systems feature profile. Based on the outcomes of studying already implemented systems, the first constructive step included the conceptualization, exemplification, and classification of their (physical implementation independent) distinguishing characteristics, called **paradigmatic system features**. Here the word 'paradigmatic' is used to emphasize that only those quantitatively or qualitatively assessable (measurable or computable) characteristics have been considered, which are permanent as long as the underpinning system paradigm (ontological model) is permanent, and change only when it changes.

For Pourtalebi et al. (2014), the point of departure was that systems have both paradigmatic features and manifestation features. The corresponding abstraction levels are shown in Figure 2.3. Paradigmatic system features (PSF) can be both logically-based and physically-based abstractions of a system or a dominant part thereof, as a whole. In other words, they can be derived by abstracting or generalizing from the various sets of manifestation features of the instance systems. These features are related to the purpose or disposition of a system and either differentiate it from other comparable systems, or make it paradigmatically congruent with them. By this process, a kind of system genotypes are actually established.

Many survey papers have interpreted system features starting out from the conceptual map of CPSs, first presented by the Cyber Physical Systems Organization (CPSO, 2009). For this reason, I could observe a somewhat blurry picture again. For instance, Togay (2014) regarded: (i) tightly integrated, (ii) heterogeneously networked, (iii) multi-aspect adaptability, (iv) automation capability, (v) non-functional requirements, (vi) distributed architecting, (vii) multidisciplinary engineering, (viii) limited resources, (ix) time awareness, (x) general dependability, (xi) predictability and determinism, and (xii) risk of casualties as paradigmatic features. Talsania et al. (2017) identified: (i) closely integrated, (ii) cyber capability in every physical component, (iii) networked at multiple scales, (iv) temporal and spatial complexity, (v) dynamic reorganization and reconfiguration, (vi) high degree automation based on close control loops, and (vii) dependable and certified operation as features - most of which are associated with cyber-physical computations. With a view to s*CPSs, Horváth (2021) extended these with: (i) multi-level cooperative openness, (ii) system-level reasoning and learning capabilities, (iii) system operation in dynamic contexts, (iv) semantic, pragmatic and apobetic interactions, (v) self-supervised planning and adaptation, and (vi) ensuring multi-aspect dependability.

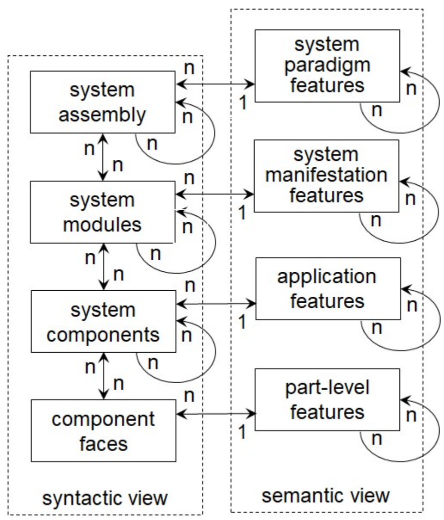

Figure 2.3: Interpretation of the features of system constituents

The paper by (Horváth and Gerritsen, 2012) presented altogether 16 paradigmatic system features that were deemed to have a selective nature at the time of doing the background research.

F_1 CPSs are designed and implemented to support human activities and well-being by distributed cooperative problem solving, in harmony with the techno-socio-economic environment,

F_2 CPSs are functionally and structurally open systems, with blurred overall system boundaries,

F_3 CPSs have the capability to change their boundaries and behavior dynamically, and to reorganize and reconfigure their internal structure,

F_4 CPSs consist of a digital cyber-part and an analog physical-part, which are supposed to work together towards a high-level functional and structural synergy,

F_5 CPSs are articulated and heterogeneous, and are constructed of very diverse sets of components, which can enter and leave the collective at any time, and may encounter other systems with similar or conflicting objectives,

F_6 CPSs, as well as their components, manifest on various extreme spatial scales (from intercontinental to nano-scales) and temporal ranges (from instantaneous to quasi-infinite), and beyond,

F_7 components are hybrid structures, encapsulating various (spatial) compositions of physical (material) entities and embedded cyber (software and knowledge) entities that provide real-time information processing capability,

F_8 components have either predefined or ad-hoc functional connections, or both, with other components at multiple levels,

F_9 components may operate according to different problem solving strategies (plans) towards achieving the overall objective of the system,

F_{10} components are knowledge-intensive and able to handle both built-in formal knowledge, the knowledge obtained by their sensors, and the knowledge generated by their reasoning and learning mechanisms,

F_{11} components are able to make situated decisions and strive for automated problem solving by gathering descriptive information and applying context-dependent causal and procedural reasoning,

F_{12} components are able to memorize and learn from history and situations in an unsupervised manner and to specialize themselves based on smart software agents and emergent synthetic intellect,

F_{13} components are able to reorganize themselves in response to an unpredictable (emergent) system state or environmental circumstances, as well as to execute non-planned functional interactions and to act proactively,

F_{14} overall decision-making is distributed over a large number of components, and is based on the reflexive interactions among the components and on multi-criteria analysis (optimization),

F_{15} different sophisticated strategies are applied in order to manage resources and maintain security, integrity and reliability of the components and the CPSs as a whole,

F_{16} next-generation (molecular and bio-computing-based) CPSs are supposed to have some level of reproductive intelligence.

Since every feature is formulated as a partial narrative definition, the above list itself is an **extendable meta-definition**. This extendibility is useful from the viewpoint of adaptation to new developments. However, no matter how complete the lists of hypothesized paradigmatic system features are, they do not go beyond what is eventually offered by proper narrative definitions. They fail to consider the fact that any particular paradigmatic feature characterizing a family of systems may be strongly or only vaguely,

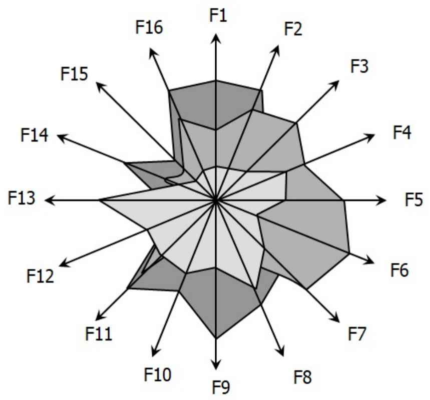

Figure 2.4: Imaginary feature profiles of CPSs

fully or partially, or explicitly or implicitly pertinent (valid) for a particular system. For instance, though it stands in general, not every CPS needs to be open, adaptive or cognizant. A significant part of the paradigmatic features discussed above imply specific operations to be achieved by intellectualization. In order to capture the set of PSFs of a particular instance CPSs or a family of such systems, and to proportionally quantify the relevance of individual PSFs for a system, the concept of **paradigmatic feature profile** (PFP) has been introduced. With regard to the individual PSFs, it is a unit normalized (measurable) indicator diagram. An imaginary graphical representation is visualized in Figure 2.4. A PFP diagram can be used not only for a *posteriori* PSF analysis of multiple systems, but also in *a priory* conceptualization of the PSF of a system at the time of starting its development. Further research is needed to find general methods for: (i) defining the thresholds and ceiling quantities/qualities of the indicators of a paradigmatic system feature to qualify as such, and (ii) specifying the threshold and ceiling values for each feature in the case of a particular CPS.

The combined observational and rational specification of paradigmatic system features cannot be separated from human comprehension and the scientific/technological progression. There is no theory on the horizon that could provide an absolutely objective view and formulation. Paradigmatic system features usually change when novel technological possibilities and engineering principles are operationalized. They may be seen also from different viewpoints. These raise the issue of inherence of system features between subsequent generations. In a subsequent generation of CPSs, the system features of the preceding generation may be preserved, partly modified, or completely replaced by novel ones. This process is referred to as the **onward transfer of paradigmatic system features**. Pourtalebi and Horváth (2016) argued that a robust taxonomy of paradigmatic system features can be derived by considering a limited set of generic system properties. Their starting point was that all systems are characterized by four foundational properties, which can be used for identification: (i) the identity property, (ii) the unity property, (iii) the hierarchy property, and (iv) the equilibrium property (Vegetti et al., 2021).

The **identity property** eventually concerns the (computational) intellect of operation and declares that changes and conversions in material, energy, information, and intellect processes are the essence of the operation of all CPSs. This property allows identifying (i) transforming, (ii) informing, (iii) reasoning, and (iv) hybrid types of CPSs (which produce outputs in the physical, digital, and cognitive realms or in all of these). The **unity property** designates existence in space and time and asserts that a system is a unique collection of components, which are more strongly bound to each other than to the surrounding environment. This property makes the boundary and the foundational rules of a CPS local. The **hierarchy property** is about the architecture and internal connectivity. It states that a system is a hierarchical composition of stable constituents that are architecturally interlocked and operationally synergetic. This property makes a CPS "more than the sum of its parts" and eventually leads to compositionality through composability. The **equilibrium property** is about the offerings (the benefits) of CPSs. It claims that outputs of a system are in balance with its inputs (that is, the laws of conservation of mass and energy, and laws of distribution of information and knowledge apply). This property is associated with the operational embedding of a CPS in its environment as well as with the use of signals/data and resources. In addition, it also appears related to efficiency, problem solving, supervision and automation issues. Apparently, these four foundational properties designate the proper dimensions of deriving paradigmatic system features. However, while it can be done relatively easily in retrospective analyses of past implementations of CPSs, it is challenging in the case of predictive investigation of future generations.

2.6. Smartness as a realistic goal of intellectualization

As discussed above, natural intelligence, computational intelligence, and systems intelligence are still largely open issues and topics for the '*terra quaestuosa*'. There is an ongoing and unsettled debate about the very essence of system intelligence, beyond its current computational implementations and applications. This debate has been out there for almost 60 years, or even more, but the light cannot be seen yet at the end of the tunnel. Hollnagel (1993) posited that “(T)here is one main problem in defining what an intelligent system is: there are no good definitions of intelligence. There is generally more agreement on the behaviors referred to by the term (the phenomenology of intelligence) than on how they can be interpreted or categorised”. Nevertheless, there are some constructive definitions, for instance that of the American Institute of Aeronautics and Astronautics, which interprets the necessary properties of system intelligence as: (i) learning (acquiring new behaviors based on past experience), (ii) adaptability (adapting the response to changing environment or internal condition, (iii) robustness (consistency of response across a broad set of circumstances), (iv) abstraction (turning data into information and then into actionable knowledge and wisdom), and (v) extrapolation (acting reasonably when faced with not previously experienced circumstances).

Some 25 years ago, Shaw (1998) argued that “(T)he ultimate goal of intelligent system design is the creation of autonomous systems which can perform complex control tasks under all operating conditions of a plant or process, even in the presence of failures, without human intervention or supervision. In this case, it will be sufficient to tell the system what to do but not how to perform the task given”. The lack of generic definitions of human intelligence, the uncertainties of defining what system intelligence ultimately means, and the current stage of the technologies enabling intellectualization of engineered systems forced us, i.e., the researchers at the Section for Cyber-Physical System Designs of the Faculty of Industrial Design Engineering, at the Delft university of Technology, to follow a less ambitious goal. This aimed at the development of theoretical fundamentals, computational enablers, and practical application cases for s*CPSs.

The line of reasoning of the conceptualization process has interrelated the three phases of architectural, functional, and computational specification. It originated from the frequently cited **5C layered architectural model** of CPSs (Figure 2.5). This model, which

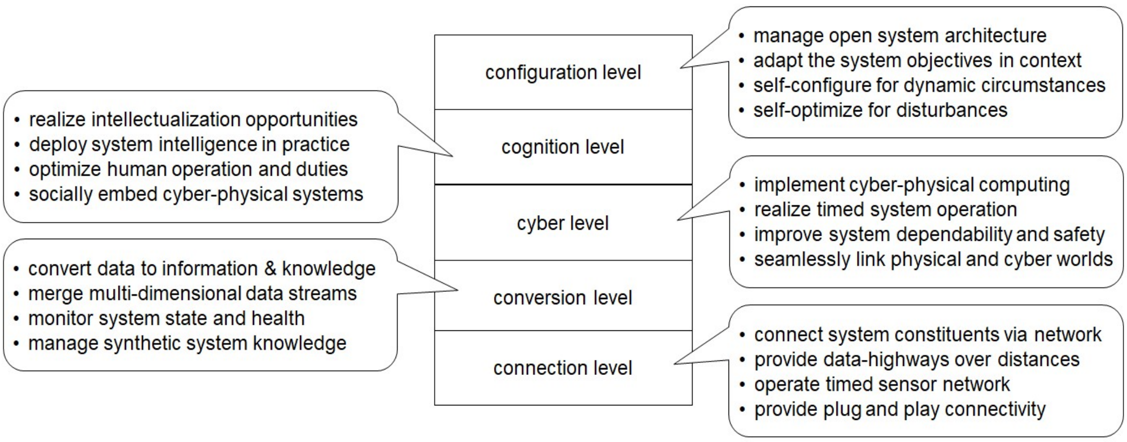

Figure 2.5: The 5C layered architectural framework

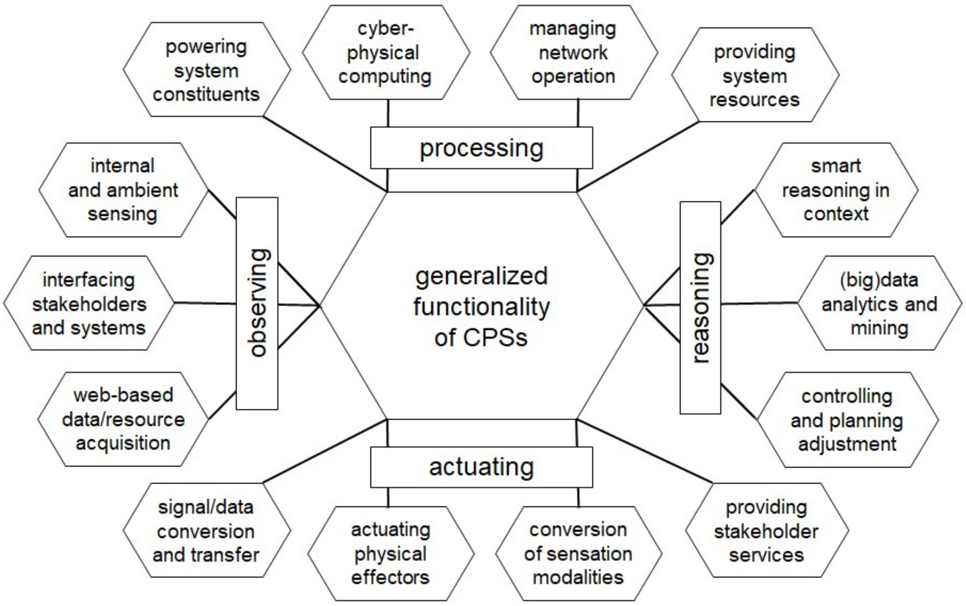

Figure 2.6: Generic functionality of i*CPSs

is actually a generic framework by its nature, includes a cognitive layer that is dedicated to supporting implementation and utilization of intellectualization (Pivoto et al., 2021). However, practical operationalization of the cognitive level constituent of the framework needs specific functionality. In practice, it meant revisiting the **generalized functional structure** of CPSs and extending it with a computational reasoning orientated subset of functions. As shown in Figure 2.6, three main functions, namely (i) smart inferring in context, (ii) massive data analytics and mining, and (iii) planning and controlling adjustment have been defined. In the stage of specification of the computational and resources, the computational mechanisms and algorithms have been determined and exemplified.

Insisting on the aforementioned pragmatic treatment, the ideal of overall systems intelligence was dropped as a target of the research activities. Instead, **system intellect-based smartness** has been used both as a starting point and a finishing point of the then ongoing research inquiries and development attempts concerning intellectualization. To sharply demarcate it from the concept of intelligence, smartness has been defined as a specific, localized, and purposeful ability to know, to do, and to make. As a reasonable goal of intellectualization of CPSs, equipping them with sufficient smartness for real-life problem solving has been targeted. In my view, smartness also implies that decisions are made with the involvement of humans in the creative, transformative, or informative operations, rather than solely by computational reasoning mechanisms of a system (Schirner et al., 2013). Eventually, smartness has been regarded as a measure of the ability to act efficiently in emergent situations based on innate or acquired competences, knowledge, and information.

There has been a reasoning model proposed to capture the essence of system smartness (Figure 2.7). This has been not only published, but also operationalized in various projects with the involvement of staff members and Ph.D. students. This reasoning model assumes that smartness concurrently manifests in a (computational) **internal operation space** (IOS) and an (observational) **external operation space** (EOS). The IOS is populated by the problem-solving core-functions enabling smart operation of a system. These functions include: (i) multi-source sensing and monitoring, (ii) awareness building of the state of the system, environment, and problem, (iii) ampliative reasoning based on system knowledge,

(iv) strategic leaning from processes and results, (v) assessing and planning the optimal operation, (vi) run-time adaptation and validation, and (vii) actuating the generative, transformative, and informative processes. In Figure 2.7, the spiraling curve depicts system smartness as an evolving phenomenon enabled by evolving problem-solving mechanisms and algorithms.

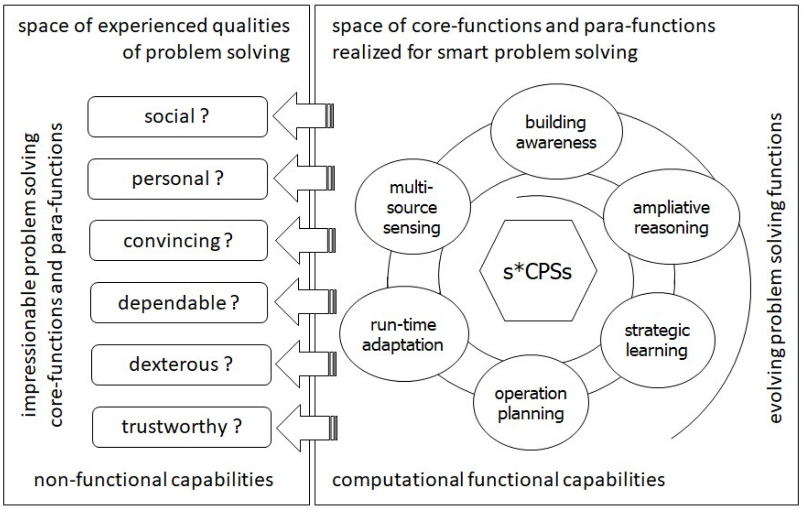

Figure 2.7: The dual-space of manifestation of system smartness

It is an important insight from behavioral science that smartness arises when problem-solving intelligence meets application and observational contexts. This demarcates the IOS and EOS spaces in the overall manifestation space of system smartness. System smartness is also a judgment from a non-functional (experiential) perspective. The EOS space is where the smart problem solving qualities are experienced and assessed. This space also accommodates the reflections on the impacts and effects that the operation of a smart system evokes. This judgment may come from any actor of the application environment, including humans and other systems. The experiences of the system stakeholders reflect the observed 'goodness' (e.g., efficiency) of the applied problem solving functions/methods, as well as the impressions raised by the operationalization of para-functions. A non-subjective judgment of the experienced system's performance requires objective and proper measures. These are not available in the literature yet. Nevertheless, we regarded: (i) ingenuity (inventiveness in human context), (ii) dexterity (agility), (iii) convincingness (proficiency), and (iv) dependability (reliability) as observable non-functional characteristics in the context of impressionable problem solving performance of s*CPSs [1]. This reasoning is based on the analogy of experiencing smartness in human problem solving. The mentioned characteristics interconnect (i) the observed sophistication of system functionalities, (ii) the experienced quality of smart problem solving, and (iii) the subjective assessment of the stakeholders of the system.

Due to its complexity, only parts of the above-discussed reasoning model have been implemented and tested in Ph.D. research projects. For instance, awareness building, ampliative reasoning, and context dependent informing have been implemented in the promotion research of Li, Y. (2019). Ampliative reasoning, strategic learning, operation planning, and effector actuation was implemented in the work of Li, C. (2016). Multi-source sensing, strategic learning, operational scenario planning, and run-time adaptation were studied by the promotion research of Ruiz-Arenas, S. (2018). The studies of Abou Eddahab, F.-Z. (2020) connected big data processing with semantic information generation and exemplified smart functions for next-generation intellectualized systems. The research of Tepjit, S. (2022) can be placed on the border between the space of the smart problem solving functions and the space of human experience and interaction with a CPS. It addressed the issues of system-assisted development of reasoning mechanisms for smart CPSs. The promotion research of van Doorn, E. (2018) dealt with information engineering for supporting situation awareness in a specific context. Pourtalebi, S. (2017) contributed to the development of the concept of the system paradigmatic features as well as to its operationalization in the design of genotypes, phenotypes, and prototypes as transitive models of CPSs. The work of Du Bois, E. (2013) contributed a stakeholder-centered, designerly methodology to pattern-based software development for smart systems. The synthesis of the novel findings revealed that it is not enough for smart systems to be able to purposefully apply computational reasoning and to solve complicated problems. They should also be able to raise the impression in the stakeholders and the whole environment that both their behavior and the solution are smart. This can be facilitated by accompanying para-functional abilities.

2.7. Cyber-physical systems for everyone

The title of this section does not want to suggest that every human being should possess one or more complicated and expensive CPSs in the near future. It is more about the opportunity of orientation of the objectives of development and application **towards non-industrial utilization**. This idea popped up more than a decade ago, and smart cars, grass cutters, kitchen assistants, bathroom cleaners, stroke rehabilitators, autonomous learning aids, etc. are proving its feasibility and usefulness. Everyone may and must benefit from the context-dependently tailored services of CPSs (Suh et al., 2014). Talsania et al. (2017) argued that CPSs are an opportunity for humans to get closer to nature by means of the cyber world or ubiquitous computation. Research on socially embedded CPSs as well as on cyber-physical-social systems has also been intensified (Zhou et al., 2019).

In the previous decade, CPSs have mainly been conceptualized and implemented as industrial enablers, integrating industrial Internet of things systems, process management and monitoring systems, computer aided design and manufacturing systems, and data/information management and deposition systems. As argued above, next generation CPSs will manifest not only in the form of large-scale complex and plant-type industrial systems of systems, but also in socialized and personalized interconnected systems formed by interoperating smart actor nodes of interconnected systems (Wan et al., 2011). As such, they will be able to (i) deeply penetrate into real life processes, (ii) collect massive amounts of data, (iii) develop operation plans run time, (iv) optimize their operations and services, and (v) adapt themselves to dynamically varying environmental circumstances or operational states, in addition to (vi) guaranteeing dependability, safety, and privacy.

The above-mentioned functionalities are becoming the distinguishing characteristics of smart CPSs for everyday applications. In the case of such systems, it can be expected that the focus of design will shift from a functionality orientation to an affordance orientation. The largely autonomous systems will decide on what affordances can be useful in a particular application context and how these affordances can be transformed into context-specific operations. Yamanobe et al. (2017) posited that the **concept of affordance** could be a key to realize human-like advanced manipulation intelligence, especially in unknown situations. In fact, only our creative imagination and the technological and economic opportunities may constrain the varieties of systems that can be developed for the exploitation of functional and technological affordances in various non-industrial applications. The growing number of related publications evidences the interest in developing and studying such application systems.

In everyday applications, smart CPSs need novel (i) computational mechanisms to recognize operational affordances in a concrete problem solving process, (ii) abilities to adapt themselves according to a possible realization of a particular operational affordance, and (iii) resource management and additional/different resources for operationalization. We have investigated the computational implementation opportunities of these and published the results (Li et al., 2014). The proposed procedure includes: (i) identification of states of the system and the environment based on run-time sensed signals/data, (ii) recognition of the changes in the state of the system and the environment as events, (iii) inferring about situations characterizing the system and the environment based on the whole of the temporary events, (iv) building awareness through monitoring the situations, (v) assessment of the performance of the system with regards to its operational/servicing objectives, (vi) devising alternative performance enhancement options by prognostic reasoning, (vii) adaptation of the constituents and the whole system according to the best enhancement option, (viii) devising and scheduling the implied interventions in the system and the environment, and (ix) actuating the concerned effectors and controls. These activities are driven by the principles of the system-in-the-loop (Hartmann et al., 2017) and the environment-in-the-loop (Falkenberg et al., 2018) approaches, rather than by the principles of the human-in-the-loop approach.

Chapter 3

3. Trend of progression of intellectualized cyber-physical systems

3.1. Research objectives and approach

The literature of CPSs mirrors a perpetual shift in terms of the scientific fundamentals and the derived concepts used in this family of systems. The situation characterizing the last decade was described by the National Academies of Sciences, Engineering, and Medicine (NASEM, 2016) as "today's practice of CPS system design and implementation is often ad hoc, not taking advantage of even the limited theory that exists today, and unable to support the level of complexity, scalability, security, safety, interoperability, and flexible design and operation that will be required to meet future needs". The researchers of different institutions or platforms have provided **differing interpretations** of this shift and have formulated different visions of its destination.

It is not incidental to find publications whose authors underpin their work with theories and methodologies that belong more closely to other fields of system engineering, such as embedded systems, application software, wireless networks, or digital twins. In other words, dissimilar system categories have been referred to as CPSs or constituents of these. This phenomenon can be explained by the fact that the different emerging categories of CPSs have not been identified uniquely. Socialized robotic systems, agent-based actor systems, complex adaptive systems, Internet-of-things systems, and the like, are put into the same functional basket and their names are often used interchangeably. Furthermore, the distinguishing functionalities, capabilities, and features of the successive implementations of CPSs have not been described in a consistent manner. The differences between smart, aware, responsive, intelligent, context-driven, proactive, autonomous, etc., CPSs are not made explicit. Certain early implementations of CPSs have already been identified as intelligent systems, while they are just advanced implementations according to the criteria of other researchers (Claudi et al., 2013).

The goal of this research was to devise a **reasoning model** considering the generation-shift hypothesis as a theoretical basis. Acknowledging the usefulness of parsimony, our intention was to find a minimal number of aspects which a robust categorization may rest on. The specific goal of the investigations was to logically demarcate various kinds of CPSs based on these aspects and the observed or supposable paradigmatic features and to identify non-isomorphic generations (i.e. distinct taxonomical categories that cannot be mapped onto each other). The concept of 'system generations' and the proposed reasoning model will be further elucidated in sections 3.2 and 3.3, respectively. Sections 3.4 - 3.8 discuss the main characteristics of the zeroth, first, second, third, and fourth generation systems, while section 3.9 addresses some open issues deemed important for near-future progression. The contents of this chapter have been compiled from the following peer-reviewed publications: H9, H10, and H11 (see Appendix A.1.1).

3.2. Determinants of progression

The aforementioned findings of our preceding research have raised a number of research questions such as: 'What characterizes the progression of CPSs?' 'What differentiates past, current and future CPSs?' and 'What milestones are in the process of progression?' Answering the first question boils down to identification of the determinants of the progression and how they actually influence the manifestation of CPSs. Obviously, the overall progression is framed by the industrial revolutions, which combine the results of (i) advancement of science, (ii) development of technologies, (iii) changes in industrial, commercial, societal demands, as well as the effects of (iv) alteration of human thinking, and (v) the affordances and limitations of the environments. The associated theories are rarely specified in enough detail to enable systematic evaluation of their assumptions, mechanisms, factors, and outcome. The only exception is technological development, in particular computational technology development, where many deep-going trend analyses, quantitative technology forecasts, model-based predictions, and convergence assumptions are available. This applies to the development of artificial intelligence and intellectualization of systems (Brundage, 2015). Nevertheless, at the beginning of the 2010s when we completed our explorative research, we could not find any model that would have characterized the progression of CPSs. Typically, static categorizations are available, like the one proposed by Thekkilakattil and Dodig-Crnkovic (2015), which systematized the types of CPSs as: (i) automatic, (ii) semi-automatic, (iii) semi-autonomous, and (iv) autonomous systems.

Towards such a model, we started out with **two general hypotheses**. We hypothesized that capturing anthropic potentials could lead to a relevant basis for modeling the progression of systems. If systems are working in the physical (mechanical, biological, physiological) domain as well as in the cyber (digital, virtual, logical) domain, then they must be able to organize the physicality and to manage the intellectuality simultaneously. We contemplated the analogies of the body and brain. In this line of reasoning, we have identified self-organization capability and self-intelligence capability as determinants of the progression of CPSs (Figure 3.1). We have implicitly assumed that their interaction at various progression levels is what actually defines the manifestation of systems. This aligns with the principles of cybernetics, which identifies first, second, and third order regulatory systems. Our analysis has explored that it is not necessary to consider self-autonomy as a determinant of progression, since it is an outcome of possessing the respective higher levels of self-intelligence and self-organization (whereas it cannot go beyond the respective forms of automatism or automation at lower levels of these capabilities).

3.3. Proposed reasoning model of progression

Our second hypothesis also rests on the natural world, more specifically, on the biological realm. In this realm, succession is typically characterized by the appearance of a new generation and by the genetic differences between the subsequent generations. This analogy has already been applied and tested in genetic computing. Our assumption has been that the historical development of CPSs can be properly described by the concept of **system generations**, like in the natural word. It has also been postulated that the emergence of generations can be brought into direct relationship with the shift of system paradigms. A 'generation' of CPSs has been defined as the total of differentiating system concepts, paradigmatic features, architectural principles, functional abilities, technological implementations, and offered services. Some or all paradigmatic features of the

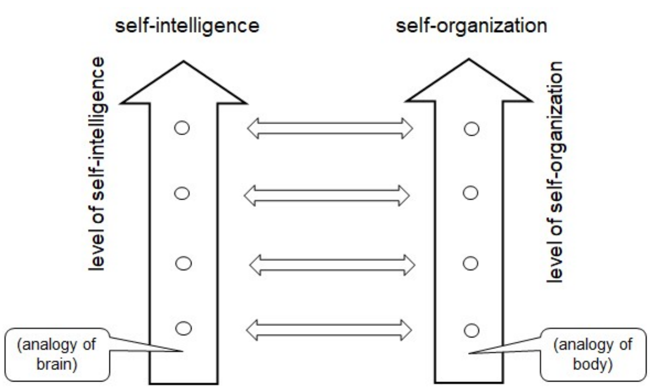

Figure 3.1: Assumed determinants of the progression of CPSs

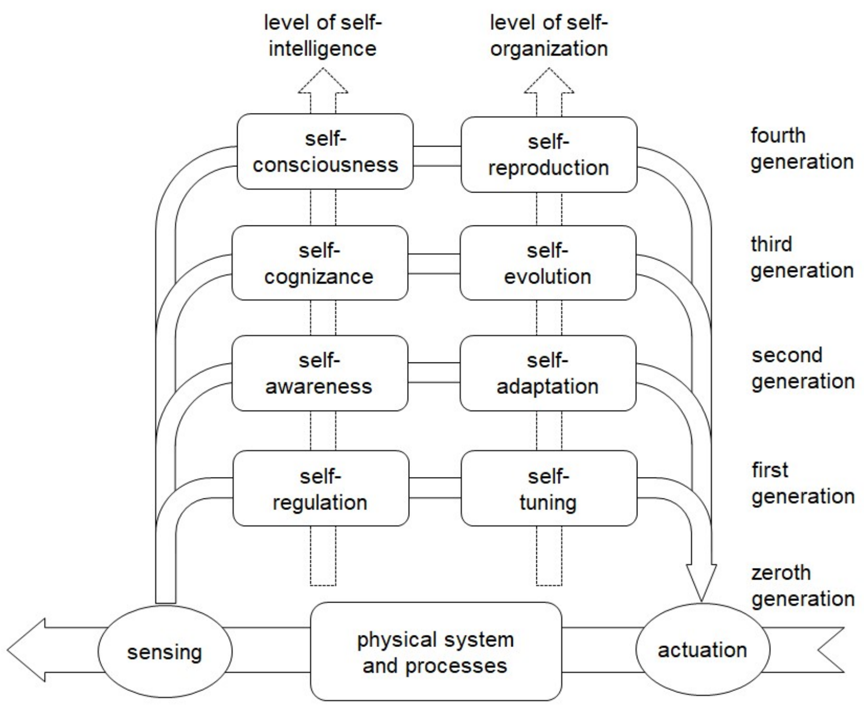

Figure 3.2: The conceptual advancement model

antecedent systems may be preserved in the paradigmatic features set of a particular generation, but they can also be replaced or supplemented by specific novel features. This interpretation of subsequent generations assumes aggregation and substitution mechanisms rather than abrupt changes behind the transitions of generations. This kind of progressions obey the principle of the 'shift of paradigms'. As a structural term, 'generation of CPSs' means a 'technological and engineering cohort' of different individual manifestations of systems. Like biological generations, which may exist in similar or different periods of time, generations of CPSs may also be coexisting or having overlapping existence. This is another aspect of the 'shift of paradigms'.

Based on the above considerations, a **conceptual advancement model** (CAM) has been conceptualized, a graphical representation of which is shown in Figure 3.2. This is both a conceptual framework and an explanatory scheme. As a framework, it proposes a logical arrangement of the successive implementations of CPSs according to their growing intellectual and organizational abilities. As an explanatory scheme, it makes the envisaged stages of progression perceptible through a limited number of major generations. These have been named (and indicated by acronyms) such as zeroth (0G-CPSs), first (1G-CPSs), second (2G-CPSs), third (3G-CPSs), and fourth (4G-CPSs) generation systems, respectively. The detailed specification and description of the five generations of CPSs happened based on underpinning facts and conjectures published in the literature in the mid-2010s.

Without going into the technical details which are presented in the next sessions, the **differences between the generations** can be highlighted as follows (Figure 3.2). The 0G-CPSs include look-alike engineering systems and partial implementations of CPSs. The 1G-CPSs include systems with self-regulation and self-tuning capabilities, while the 2G-CPSs have the capability to operationalize self-awareness and self-adaptation. The 3G-CPSs are equipped with the capabilities of self-cognizance and self-evolution. Only the 4G-CPSs are supposed to achieve self-consciousness and self-reproduction in the form of creating a system of systems. 4G-CPSs will be based on an integrated model of intelligence, which, however, needs to be implemented at full-scale. Addressing both architecture (structural) and operation (intellectualization) issues, the next sections further analyze and discuss the paradigmatic characteristics of these generations.

3.4. Alternatives of self-organization

Self-organization is realized differently by the different generations of CPSs (Figure 3.3). This simultaneously concerns the interrelated functionality/operations and architecture/configuration of the systems. Every point of the **initial system space** (ISS), the **extended system space** (ESS), and the **reproduced system space** (RSS) represents a particular functional and architectural organization of the system. Below, I am going to interpret the alternatives of self-organization as it influences the operation of the system. 1G-CPSs always work within the **designed operation space** (DOS) and remain inside the ISS, which is enabled by the predefined resources possessed by the system (Figure 3.3.a). An exception can be malfunctioning of the system. The DOS can be anywhere on the ISS, assuming that each of the selected individual operational states is feasible (and meaningful). The system can self-tune itself by selecting alternative operation states, and find a relative optimal operation state by setting all system parameters accordingly. It may happen that the operation of the system remains sub-optimal due to the fact that it cannot extend the space of possible operations due to the restrictions in terms of the system parameter values, or the limitations emerging due to resource constraints.

2G-CPSs may extend the ISS in order to achieve a new optimal objective. Based on its adaptation abilities, it can create an extended system space (Figure 3.3.b). Starting from the DOS, it may search for an **optimum operation space** (OOS) not only on those parts of the ISS, which are enclosed by the ESS, but also on the whole of the ESS. Certain parts of the ISS may lose their relevance to searching for an OOS, due to the system parameters which become obsolete as a result of the adaptation. It means that, in some cases, extension of the ISS may actually mean (i) expansion, (ii) contraction, and (iii) a combination of space expansion and contraction. The continuity of the ESSs is maintained, that is, no new system parameters are introduced. The multiplicity of the ESSs makes adaptive self-organization possible.

3G-CPSs may sub-sequentially extend their ISS to various ESSs multiple times and according to various contexts, and may dynamically search for the OOS on any one of these

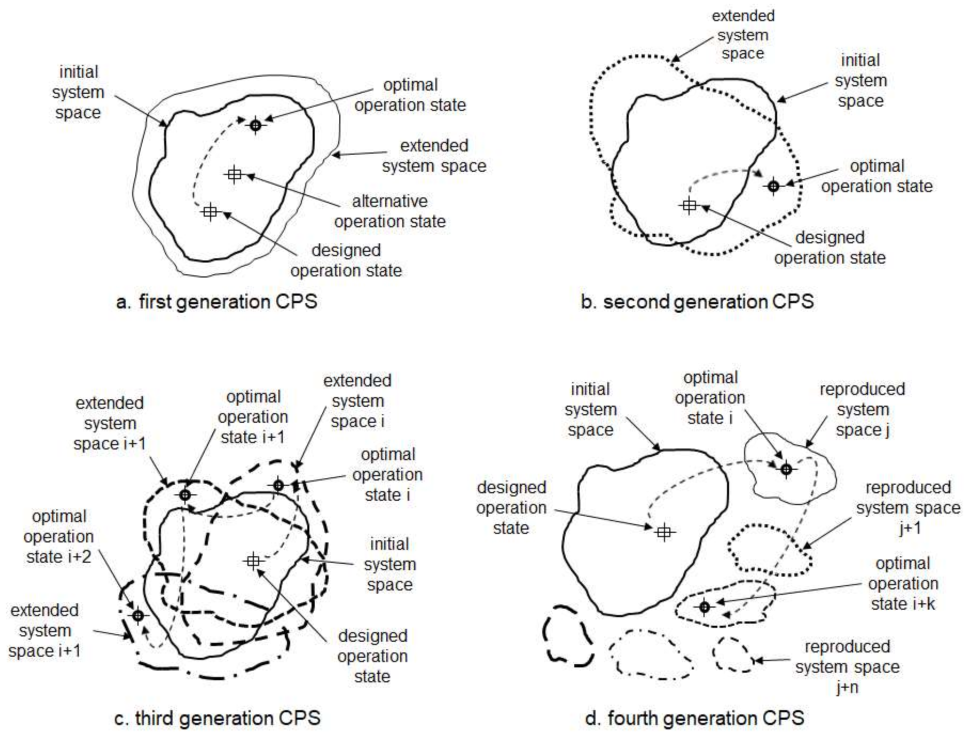

Figure 3.3: Alternatives of self-organization

EESs (Figure 3.3.c). That is, the OOS can be found on any one of the ESSs. As before, certain parts of the ISS and some ESSs may lose their relevance to searching for an OOS. The continuity of the ESSs is not maintained, that is, new system parameters are introduced. This process allows extensive search and learning processes. The multiplicity of the ESSs, the new system parameters, and the learning abilities lend themselves to an evolutionary self-organization.

4G-CPSs may generate multiple RSSs, which may be completely disconnected (Figure 3.3.d). Each of the RSSs represents different operational objectives and operational system parameters. They are managed by sophisticated reproduction strategies. The self-organization process of the systems starts with the establishment of variant(s) of the system that correspond to the RSS(s), and then the search for an OOS is executed. These may be repeated multiple times, according to some dynamic objectives. The reproduced EESs may be decentralized and distributed replicas of the ISS. The above overview shows that the various approaches to self-organization create CPS generations of rather different behaviors.

In the last two decades, both designing for adaptation and designing for self-adaptation have become protruding design methodological issues in application contexts. The fact of the matter is this is a hot research topic in the literature related to software systems. It is also influenced by the high variance of types and applications of i*CPSs. Recently, system adaptation has been identified as a key technology towards autonomous driving (Haböck et al., 2016). In addition to traffic management, energy provisioning, and manufacturing environments, adaptive systems have been penetrating into the domain of medical systems too (Abbod et al., 2002). Brown (2006) elaborated on the application of complex adaptive systems theory to clinical practice in rehabilitation. At the same time, among others, Weyns et al. (2022) considered the opportunities of designing self-organization-based evolutionary software systems.

3.5. Features of zeroth generation cyber-physical systems

As the literature shows, the majority of the industrial 0G-CPSs are the result of incremental automation. This generation of quasi-CPSs includes system implementations whose architecture and operation resemble those of the truly ‘cyber-physical’ systems, but partiality and incompleteness can be observed commonly with regards to the full functional spectrum. Typical 0G-CPSs are such as embedded hardware/software systems, software-integrated plant systems, Internet of things systems, complicated production systems, medical monitoring systems, machine assembly robots in the automobile industry, and so forth (Leitão et al., 2016). Many such systems have been produced by including multiple sensing-computing-actuating loops. This has been termed as **cyber-physical augmentation** (CPA). They typically consist of one or more traditional ‘plant-type’ subsystems or monolithic artifacts, such as advanced robots, that form the physical subsystems.

With a view to their architectural and operational characteristics, 0G-CPSs are functionally and architecturally closed systems and do not lend themselves to any run-time variation. They are controlled by predefined (preprogrammed) closed loop control and optimization subsystems. The control is typically model-based, generated based on the principles of classical control theory. In many cases, it involves networked control. The applied classical control theory assumes that the controller continually or periodically extracts continuous or discrete-time signals concerning the state of the physical processes and the system as a whole, and continually or periodically actuates the effectors in order to achieve the objectives of the system’s operation. The primary objective of using a closed-loop control subsystem is achieving accuracy, stability, and reliability. To handle the realities of software and networks, adapting control theory is also applied, but it is an unsolved issue for many systems (Hou and Wang, 2013).

3.6. Features of first generation cyber-physical systems

Shown in Figure 3.2, the distinguishing paradigmatic features of 1G-CPSs are **self-regulation** and **self-tuning**. Feedback-based self-regulation and self-tuning characterize the lowest level of proactive and deliberate system smartness and adaptability, respectively. Typically, 1G-CPSs are closed and software-integrated systems. A set or network of sensors measures the parameters of physical processes, and the measurements are processed in the cyber subsystems, which provide data for driving the actuators that affect the physical processes. These systems include algorithms and software components that collect the sensor data and react on them by issuing control signals via the actuators to the physical effector components. With regard to the dynamics of the operational conditions and system-level interactions, the control subsystem needs to cope with the dynamic phenomena of software and networks, such as: (i) timing jitter in communications and computation, (ii) packet losses in networks, or (iii) resource contention having profound effects on the performance and stability of the physical subsystems. However, the control system applied in 1G-CPSs is neither adaptive, nor predictive. Typically, 1G-CPSs can handle only anticipated changes, that is, changes that occur within the operational domain for which the system has been built (Wu and Kaiser, 2012.

The main architectural and operational characteristics of 1G-CPSs are as follows. They are usually designed as independent (self-reliant) systems. However, some of them make the first step from simple independent systems towards interlinked system(s) of systems. The term ‘self-regulation’, used in Figure 3.2, refers to the ability of reaching a very limited (low level, preprogrammed, or self-acquired) of system ‘intelligence’ in given contexts (based on syntactic rules or processing probabilities). Self-regulation is exemplified by the traditional feedback type of control or learning enhanced control, which enables self-tuning. This expresses the restricted potential for functional and architectural adjustments according to different operational conditions (without any serious change in the goals, architecture, and operations). Typical is self-resilience, which usually means activating redundant components in the case of a heavy working conditions, malfunctioning or intrusion (for instance, think of a lorry, which lets a third rear wheel-pair down when overloaded and lifts it up when no extraneous payload is there, but can reduce rolling friction this way. Resilience is the persistence of service delivery that can justifiably be trusted when facing changes, i.e., the persistence of dependability when facing changes. Resilient 1G-CPSs are the next challenge – adding new dimensions of complexity to manage (Chaterji et al., 2019). The control strategies implemented in the cyber subsystems need to achieve self-tuning (responding to changing operational conditions). However, they usually do not have capabilities for a predictive behavior (anticipating changes in the physical, communication, and computational processes).

3.7. Features of second generation cyber-physical systems

Developers of 2G-CPSs strive after handling unanticipated changes, such as anomalies outside their operational domain, and emergence of new goals or new technologies. **Self-awareness** and **self-adaptation** have been regarded as distinguishing paradigmatic features of 2G-CPSs. Between these two, there is a functional relationship established by the operational/servicing objectives. Self-awareness of intellectualized systems is often reduced to the possession of information and inferring based on this information in different contexts. Therefore, it is often manifests as computational **self-monitoring**. Self-adaptation may be necessary due to (i) the growing complexities of system operations and servicing, (ii) the limitations of taking all possible changes in the system and the environment into consideration at the design phase, and (iii) the uncertainties/alterations of the operational objectives, environments, circumstances, and human interactions.

One of the challenging questions today's research is facing is how to get to and operationalize actionable insights through and by systems themselves (Endsley, 2018). Regarding self-awareness of 2G-CPSs, it is of importance to differentiate between extensional awareness (that includes reference to an extensive knowledge structure on the state of the system and its environment) and intentional awareness (that includes a natural consciousness based insight into the state of the system and its environment). Thus, as formulated in our research, the working definition of system awareness states that it is a computational capability and complex mechanism that: (i) captures local word related, (ii) purpose and task dependent, and (iii) context and situation influenced information construct(s), and gradually builds up (iv) a comprehensive and semantically rich model, and (v) an operational 'mimicked understanding' in a particular system. The main issue is that there are multiple stimulants (e.g. topical interest, body language, eye contact, facial expression, etc.) in the case of humans that support building awareness through back-channel feedback and nonverbal cues (Greenberg et al., 1996). These cause the real complications for modeling and computation, since they cannot be captured directly as para-functional abilities and transformed into information structures and procedures in a straightforward manner.

Self-awareness assumes a near-zero-time processing of perceptional information and knowledge, and a kind of spontaneous recognition of objects and events, even in dynamically changing situations and contexts. Context information processing by 2G-CPSs should amalgamate: (i) run-time elicitation and structuring of context information, (ii) modeling and analysis of context variability, (iii) run-time integration of context information with reasoning/decision making procedures, (iv) reconfiguration, rebinding, and dynamic-composition of changing context information, and (v) use of concept ontologies for disambiguating and discriminating context information at run-time. Low-level context can be represented with data entities and data constructs, which are extracted by filters from sensor data or generated based on computed data. A representative example of this is a CPS that has been developed for detecting and enhancing short term engagement of stroke patients in rehabilitation exercises (Li et al., 2016). Traditional context models are not sophisticated enough to qualify as the basis of relevant computational approaches. High-level contexts can be represented properly only by dedicated informational structures, which capture semantic and pragmatic relations, constraints and rules (Mylonas et al., 2009). This entails the need for more sophisticated transformers and schemes, which can nevertheless be processed in quasi-real time.

Adaptivity is the measure of the change potential that a system has in order to meet pre-set or possible (emerging) system objectives (Black et al., 2014). An **adaptive system** can provide gains if it has an excess of resources and components that can be rearranged, replaced, combined, or interchanged easily. System adaptivity depends on (i) the types and number of components, (ii) the degree of connectedness and the variety of the interfaces, as well as on (iii) the software components involved in processing dynamic contexts and in logical/semantic reasoning. Modularity has been identified as the relative measure of a system's ability to remove, add, or rearrange components at various levels of (de)composition. These factors together also enable moving towards self-evolution.

The current generation of adaptive systems are closed systems, and suffer from limitations with regards to the theoretical adaptation of functionality (modes of operation) and architecture (management of resources). 2G-CPSs take the first step towards opening the system boundaries from both architectural and operational aspects. They typically use run-time acquired data in addition to the data stored in the system's operation model (and control model) for system control. However, neither of these is a trivial problem, and the literature is incomplete and fragmented about resolving these issues. Having the aforementioned operational characteristics, 2G-CPSs bring the idea of smartly behaving engineering systems to reality. They are implemented as software- and knowledge-intensive systems, which are able to: (i) process dynamic context information, (ii) modify the architecture and operation runtime, and (iii) exploit run-time variability in achieving its objectives (Gabor et al., 2016). Not only is the 'designability' of 2G-CPSs influenced by their dynamic adaptation character, but also the opportunities for verification and validation of their self-adaptation as discussed in Chapter 4.

3.8. Features of third generation cyber-physical systems

With regards to the definitional conventions, 3G-CPSs are still less discussed (and perhaps, less studied) than the second generation of these systems. Their highest-level paradigmatic features have been defined as **self-cognizance** and **self-evolution**. It was discussed in the previous section that, in the order of mention, awareness and cognizance, represent different levels of intellectualization. Then, what is the essence of systems cognizance? To clarify this, we should start out with the phenomenon of human consciousness (subjective experience) (Signorelli, 2017). In the case of human beings, consciousness is a glorified, all-embracing impression and perception of the world around us. In a sense, consciousness means a fundamental state and informatization of the human brain. Human cognizance represents the manifestation of its specific implications in the cognitive space. Human cognizance is typically derived and distributed over many individuals working together towards a shared purpose or objective. Though it is shared by the concerned individuals, it may reflect differences in terms of their personal understanding and can be time-wise articulated. Thus, cognizance and awareness are two differently scoped insights and mindsets associated with local worlds and related reasoning. Cognizance is eventually a move towards consciousness (Demetriou et al., 2018). It assumes a changing scope from local to global comprehensiveness.

While self-awareness of CPSs is supposed to be the potential of a system to build a world model effective in given situations, self-cognizance is supposed to enable the development of multiple (but a restricted number of) models of the external world from various perspectives. Thus, self-cognizance is interpreted as the capability to capture and assess what the relationships in a given local word and environment are and what is happening in them. It should offer multiple computable models of the system and its environment from various perspectives. This is the basis of a '**pseudo-understanding**', which: (i) is less global than what may be expected from system consciousness, (ii) largely depends on the perspectives and contexts of interpretation, and (iii) may manifest just as a specific volatile computational understanding. 2G-CPSs still face limitations to replicate understanding conveyed by genuine human cognizance and implement different representative mental models even for pseudo-understanding.

In the case of 3G-CPSs, evolvability is continuous development towards extended functionality and improved performance. Current engineered systems are far from such a level of dynamic reorientation and reconfiguration while executing. The evolvability of 3G-CPSs appears as a **run-time performance criterion**, rather than as a challenge during the time of designing. According to the interpretation of Tackett et al., (2014), an evolvable system follows emerging objectives of system operation as often as needed. On the other hand, systems need proper resources to be able to evolve, which may be provided by system developers on request, or self-developed by the concerned systems. Notwithstanding these facts, the degree to which a 3G-CPS should be made evolvable while in service is a strategic choice in the time of designing. In the context of run-time optimizing a system design, evolution is interpreted as repeated controlled (a non-disruptive series of) reconfigurations. Some CPSs may specifically be designed to evolve, while others may

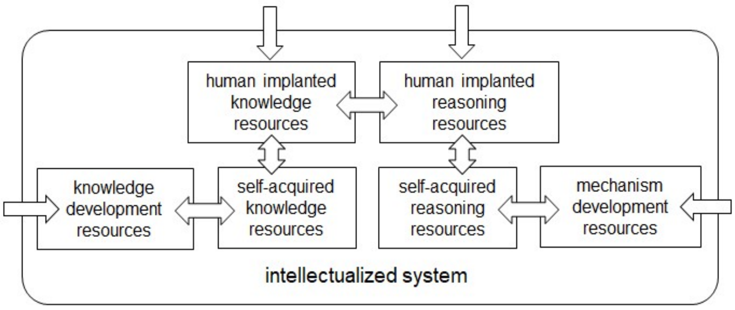

Figure 3.4: Aspects of resource management for i*CPSs

have done it serendipitously. The reason is that the (probable and possible) evolutionary transformations of a particular CPS may or may not be known during the design process, and may or may not be reversible. With a few exceptions, research on evolutionary methods for systems has been based on discrete (one-time) evolution.

The evolvability of a CPS has been defined as the **potential ability** (enthalpy or internal energy) to progress from one stable system configuration to another, and from one multi-functionality state to another in response to changes in the requirements, goals, environment and the system itself, by using system excess and modularity (Borches and Bonnema, 2008). A self-managed long term evolution of CPSs is necessitated by the emergence of new technologies and protocols/codes. The need is seconded by the fact that it is becoming prohibitively difficult to predict all future operating scenarios and features of the system and the hosting environments in the design stage of CPSs. Self-evolution is also closely connected to hardware, software, and cyberware resources and to availing these during the entire life-cycle of cognizant CPSs. This extends to (i) knowledge resources, (ii) knowledge development resources, (iii) reasoning resources, and (iv) reasoning mechanisms development resources (Figure 3.4). Combined with intense resource management, self-evolution provides an opportunity to repose the system objectives and the concepts of operations in harmony with long term technological, social and environmental changes. No quantitative model of system evolvability is published in the literature (i.e. let alone concerning a framework to successfully accommodate substantial changes in run-time). In fact, handling such changes requires evolution of the computing system. Although significant progress has been made on automating the deployment and integration of new elements, software evolution remains in essence a human-driven activity.

It is obvious that 3G-CPSs should have **excess (resources)** for enabling their self-evolution. Self-constructing software research and development offered some first solutions, but the problem is much more complex than just reusing these in the context of 3G-CPSs (Oreizy et al., 1999). In addition, there is a need also for evolvable cyberware and hardware. Both the hardware and the software have to be able to dynamically and autonomously reconfigure themselves as needed (Higuchi, 1993). This would be a dynamic evolution, resembling that of living systems. One possible approach to such a system is to use two parallel system units. A primary unit is applied for normal runtime operation of an application, while a secondary unit keeps evolving in parallel. If the performance of the secondary unit becomes better than the primary unit, they are exchanged. Thus, the unit giving the best performance at the moment is enabled. Together with evolvability, there is a need for another ingredient called recoverability (Frei et al., 1999). Towards this end, 3G-CPSs should have redundancy of resources for protecting themselves, but also sufficient excess for the intrinsic ability to recover. The system excess should also be present in the knowledge possessed by these systems. The current system designs have no such capabilities.

3.9. Features of fourth generation cyber-physical systems

Though the literature is permeated with the idea of intelligent automated systems, the fact of the matter is that this creates more confusion than clarity (Tyagi et al., 2021).

Mentioned earlier, one of the key reasons is the lack of a shared transdisciplinary understanding and generally-accepted exact definitions, which are difficult to state due to the varying interpretations of intelligence. As Hunt and Jaeggi (2013) argued, the intense research in many academic fields related to intelligence has not resulted in a commonly understandable specification of what intelligence and, in particular, of what system intelligence is. Rating intelligence of systems based on how well they can do what humans do may be not only misleading, but also illogical. I share the opinion of Erickson (2014) in that the notion of "intelligence" can be used only for comparisons within a similar group, and not within the ontologically different populations of humans and systems.

For the above reasons, a kind of 'critical soberness' should be present in talking about 4G-CPSs. In the conceptual advancement model, **self-consciousness** and **self-replication** were claimed as the highest-level distinguishing paradigmatic features. These are simultaneously direct foundations and manifestations of system intelligence. However, the abstractness of these notions has to be resolved. As conceived, system consciousness would be a fully featured replica of human consciousness, which is global and decontextualized. Actually, this demands a comprehensive and deep implementation of system intelligence (including traditional capabilities such as machine perception, situation awareness, computer vision, machine learning, etc., but also many new ones). Computational system consciousness necessitates the potential of deriving and/or maintaining a large (but still infinite) number of computer-internal world models during run-time, even in cross-context perspectives (Reggia, 2013). This is deemed to be a principal difference in comparing 3G-CPSs with 4G-CPSs. Current research has not reached further than seeking for the potential theoretical fundamentals.

One candidate for this purpose is the **information integration theory** (IIT) that primarily intends to explain the mechanisms of forming consciousness in the human brain. The work of Tononi (2012) implies, and the IIT specifically claims, that consciousness corresponds to the capacity of a system to semantically integrate information. This implies two key phenomenological properties of consciousness: (i) differentiation, i.e. the availability of distinguishing a very large number of conscious experiences, and (ii) integration, i.e. the ability to arrive at the unity of each such experience. The information engineering principles of how a conscious system can be constructed with non-neural ingredients cannot be directly derived from these. Therefore, IIT leaves questions such as 'What can be a substitute for the higher neurons of the brain, which differentiates and integrates the input from sensors, and which handles and sorts out the extraordinarily large number of possible states?' unanswered.

If we accept that human consciousness is everything that we experience and conjecture, system consciousness is everything that a system can experience and conjecture. Of course, this is rather intangible. This explains why no comprehensive computational methodologies of system consciousness exist (at least according to my best knowledge). However, one general principle is implied by the IIT, namely that a CPS should be able to distinguish a large repertoire of possible constituents of the existing reality (differentiation) without decomposing it into a collection of causally independent constituents (integration). In order to generate consciousness, 4G-CPSs should not focus on the details of the constituents, but on the stored similar 'images' of constituents and their relationships. Integration (i.e. realization of the unity of each conscious experience) assumes the availability and run-time parallel processing of a huge amount of causally effective constructs of represented information. Moreover, a similarly important feature of 4G-CPSs is the evolution of knowledge, which seems to be necessary for evolutionary self-reproduction.

It is still philosophically debated what the concept of **fully autonomous intelligent systems** means. I advocate that CPSs can be called intelligent if, and only if, they have reached the level of intelligence, autonomy, symbiosis and sociality that is comparable with that of human individuals, and more importantly, of (productive) human communities. The reasoning behind it is as follows: The ultimate forms of self-intelligence (consciousness) and self-organization (evolutionary reproduction) are not all-or-none properties. Even if the issue of resource provisioning is completely ignored (but ought not to be), it is still rather unclear how the path from evolutionary reproduction does lead to autonomous reproduction, and how it needs to be supervised (should it be needed) (Andry et al., 2001)? Consequently, in a critically sober framework of thinking, there remains nothing else but to regard these as possible long term objectives (and strategic developments), rather than immediate targets of CPS development.

3.10. Reflections and open issues

Though the proposed notion of 'system generation' is speculative, it is both logical and pragmatic with regard to the rapid development of CPSs. The model clarifies that generations of CPSs will mainly differ (i) in the level of intellectualization (i.e. self-regulation, self-awareness, self-cognizance, and self-consciousness), and (ii) in the potential alteration of functionalities (i.e. functional, architectural, and behavioral self-tuning, self-adaptation, self-evolution and self-reproduction).

- The usefulness of the proposed CAM originates in three affordances. First, from an ontological point of view, the proposed conceptual advancement model lends itself to a robust categorization of the various contemporary and subsequent implementations of CPSs. Second, it offers means to capture the research and development trends and milestones of CPSs and is useful with regard to academic education and industrial advisory practice. Third, it casts light on conceptual/terminological inconsistences and can contribute to notional clarifications and to formation of a shared vocabulary. Fourth, as a kind of 'future framing model', it can support road mapping and scenario development for both industrial and everyday applications.
- The above conceptual analysis implies that we face a rapidly growing uncertainty and knowledge gap as we move towards the 4G-CPSs. Three major sources of uncertainties are: (i) the unpredictable development of cognitive and sentiment technologies, (ii) the socially and commercially unjustified saturation by NG-CPSs, and (iii) the undecidable nature of the need for near-human intelligence, consciousness, socialization, and personalization of NG-CPSs.
- From a methodological point of view, the conceptual model suggests that system development needs to abandon the traditional (multidisciplinary) system engineering approaches and to move towards truly transdisciplinary (even supradisciplinary) approaches that consider hardware, software, cyberware, and brainware aspects simultaneously and synergistically.
- The generations-oriented thinking also raises some paradoxical issues since classical system design has a definitively targeted (prescribed and fixed) end goal, but this is not really necessary according to the CAM. A growing part of the design activities can be delegated to systems.
- From a pragmatic point of view, the reasoning model shows that we need much more research and knowledge about transferring cognition from humans to complex systems with emergent functionality and characteristics. We have to identify the issues that can be addressed and resolved in the design phase by designers, and those that are to be managed by the CPSs based on their inherent characteristics. Notwithstanding the existence and influences of these, the conceptual advancement model should be augmented, since it ignores a number of things. For instance, it captures neither the varying human relationships with the different generations of CPSs, nor the social issues associated with them (Ning et al., 2016).

Chapter 4

4. Distinguishing nature of system-level problem-solving knowledge

4.1. Research objectives and approach

This chapter presents the results of the study of the nature and essence of the **problem-solving knowledge** possessed by i*CPSs and puts it in a broader context. The motivation for the investigation came from the work of Machlup (1980, 1982, 1984), who studied humanities, science, and social science, and identified them as three distinct fields of academic learning and knowing. He used the words alpha, beta, and gamma to differentiate the bodies of knowledge associated with them, in addition to general knowledge (Figure 5.1). The conducted research also found that Gilles and Paquet (1989) had identified a fourth type of disciplinary knowledge and labeled it as the delta. This includes the specific knowledge of creative disciplines such as design, law, and economy.

As the preceding chapters touched upon it, various concepts and manifestations of system-possessed knowledge have also appeared since the time of these road-paving works. That gave the basis and the legacy of research questions concerning the existence and evolution of knowledge in intellectualized engineered systems, which are supposed to collect, infer, or extract a massive amount of knowledge based on some pre-programmed human knowledge. Problem solving and behavioral knowledge can be aggregated longitudinally (in one system over time) or transversally (over multiple systems or on a system of systems level). The guiding hypothesis of the background study was that system-generated and aggregated **synthetic knowledge** (SSK) was not covered by the aforementioned four genres of knowledge. The conducted, broadly-based literature study underpinned this claim. In fact, it represents a new genre and can be termed as epsilon-knowledge. This proposal has been extensively discussed by the peer reviewers of the papers published on this topic, but also in the public media. The contents of this chapter have been compiled from the following peer-reviewed publication: H12 and H13 (see

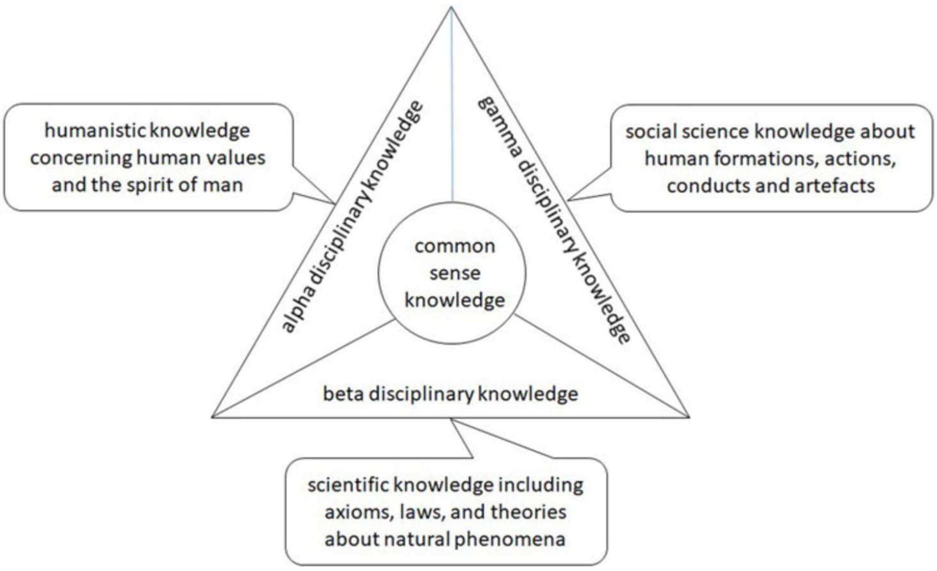

Figure 5.1: Three genres of human knowledge according to Machlup. F.

Appendix A.1.1). In order to make them comparable with delta- and epsilon-knowledge, the next sections summarize and reflect on the characteristics of alpha-, beta-, and gamma-knowledge. The chapter also provides a concise overview of the sources of SSK.

4.2. Alpha-knowledge

Representatives of humanities (also called alpha-disciplines) are such as art, archaeology, folklore, history, journalism, linguistics, literature, logic, metaphysics, music, philosophy, ethics, and theology. It is well-known that humanities conduct studies of how people express, understand, process, document, and record human experience in various modalities. These modalities include beliefs, speculations, language, literature, music, faith, creative art, performing art, history memos, to name but the most evolved ones (Agresto, 1983). The knowledge delivered by humanities has both an expressional and a sensational perspective. They together create potential connections between humans uttering experience and humans perceiving the utterance, over space, time, and mind-sets. Humanities reflect not only experience, but also diverse ideologies, cultures, heritages, traditions, and histories (Levi, 1983). Myers (1967) emphasized that there is vagueness and uncertainty about the nature of the contribution of humanities to the academic community and to the wider world of human affairs.

For the reason that the humanities are not merely academic disciplines, but also important intellectual components of societies, the knowledge associated with them is rather versatile and often carries abstract concepts (Vaziri et al., 2019). Alpha-knowledge is varied according to the disciplines of humanities. Broderick (1983) argued that humanities embrace areas of human knowledge that feature the following characteristics: (i) central concern is human beings rather than the processes of nature or the structures of society, (ii) primary focus is on the individual rather than on the group, (iii) awareness (quite self-conscious awareness) of how we know what we know, (iv) attention to moral values, whether drawn from God, man, or nature, and (v) insistence that the process of intellectual growth calls for forthright moral judgments as an equal part. Humanities work through four insubstantial mechanisms: (i) immersion, (ii) embeddedness, (iii) socialization, and (iv) reflectiveness. Typical methods of dealing with alpha-knowledge are interpretational (criticism), comparative (affects), and speculative (contemplation) in nature, though anthropology, archaeology, and jurisprudence use methods which have scientific flavor.

Alpha-knowledge is **descriptive knowledge**, which may appear in many alternative representations (text, movie, stage play, etc.). A part of humanities knowledge rests on facts and causalities, while the other part on ideas and visions, or a mixture of these (Boas, 1957). Furthermore, many disciplines of humanities reflect the duality of art and craft (Smith, 2015). It is also worth mentioning that several studies refuted the universality of knowledge in the humanities and social sciences, and that the contextuality of knowledge affirms "the social construction of knowledge" or "the enacted social knowledge", though they do not object that belief is a necessary condition for knowledge (Wray, 2007).

4.3. Beta-knowledge

Like humanities, natural science is also a philosophical category that (i) aims at understanding the natural world existing around us, (ii) produces and accumulates scrutinized knowledge as an end, and (iii) facilitates doing and making (Lakatos, 1970). After a long period of steady progression, the diversification and articulation of science accelerated in the last century (Popper, 1962). Modern science has not only empirical and rational bases, but also computational inferences and massive data streams as bases (Džeroski et al., 2007). It provides us with the epistemologically most warranted statements on the natural world, human beings, human societies, human physical constructions, and human thought constructions that can be deemed as reliable at the time of being. The scientific approach of inquiry entails the embodiment of commitment to evidence in the process. In other words, the process of inquiry is featured by the quest for the "rationally agreeable and empirically evidential" truth in the scientific disciplines. Disciplines of formal science deal with proof, while natural and applied sciences seek evidence.

Machlup (1982) called the knowledge explored and aggregated by the scientific disciplines as beta-knowledge. This genre of knowledge has a strong epistemology that clarifies not only its sources, but also the way of coming to know it, and the issues related to its justification, validation, and consolidation (French, 2007). What makes the various bodies of scientific knowledge scientific is that they are: (i) objective, (ii) rational, (iii) systematic, and (iv) universal. Objective means that: (i) phenomena are studied in an unbiased manner by alternative methods, (ii) the findings are scrutinized against empirical evidence, and (iii) the agreement of the concerned scientific community is sought to accept results as "properly justified correct belief" or "relative scientific truth". Rational expresses the condition that any research work and its results should be reproducible either in a concrete or in an abstract way, or in both. Systematic implies that the results are supposed to be generated in a systematic way to match an existing paradigm or to create a new one. Universal means that the results of scientific research are expected to be universally applicable and the conditions of usability have to be precisely stated. These, together with commitment to evidence, are often referred to as foundational criteria of scientific knowledge.

Epistemologists typically distinguish **three kinds of scientific knowledge**: (i) acquaintance knowledge, (ii) knowledge-how, and (iii) propositional knowledge (French, 2007). Beta-disciplines are concerned with propositional knowledge, which is considered historically as properly justified correct belief (Gordon, 2012). In line with the progression of studying "scientific" phenomena, scientific theories may convey (i) discovery, (ii) description, (iii) explanation, (iv) prediction, and (v) regulation (Horváth and Duhovnik, 2005). Scientific theories include statements that (i) explain all the facts in a given context, (ii) logically relate the facts based on their content, (iii) make the applicability of the theory clear, (iv) indicate what is left over of the theory, and (v) trigger hypotheses that can extend the theory to cover a broader field. In the correspondence view of science, scientific laws capture the fundamental relationships concerning phenomena and describe what happens, whereas scientific theories explain why and how a phenomenon happens.

Individuals and teams may initiate research data, hypotheses, and theories, but the consolidation of their scientific value is in the hands of the scientific community and the users. This is an element of the process of **social construction** of scientific knowledge (Mendelsohn, 1977). Though based on empirical evidence, according to the consensus view of science, beta-knowledge is tentative, theory-laden, view related, and technology dependent (Rosenberg, 2000). Science philosophy and epistemology have addressed many issues of synthesizing individualistic and social perspectives, as well as various social aspects of processes and developments related to scientific knowledge. Although there are several studies on social construction of knowledge and enacted social knowledge, little has been done with respect to "social utilization of knowledge" (Stehr, 1996). At the same time, the Internet created new opportunities and technological capabilities for structuring and refinement of raw scientific data and knowledge.

4.4. Gamma-knowledge

Among others, education, economics, environmental science, law, politics, sociology, and statistics represent present-day gamma-disciplines. These specific disciplines of social sciences are in different stages of their development – the arrow of which is pointing from an experimental character to an abstract character. Social sciences are orientated to studying the human society and social relations. Social scientific realism assumes that social realities exist independently from thinking, observation, and behavior of observers. In the middle of the last century, Mises (1942) discussed that the foundations of the modern social sciences were laid in the eighteenth century, starting with history, and that a radical change took place by picking up phenomena that belonged to political economy, human action/conduct, and social cognition/moral. He indicated the restrictions on experiencing or formal modeling and argued: “What makes natural science possible is the power to experiment; what makes social science possible is the power to grasp or to comprehend the meaning of human action”. He is among the first ones whose work identified the **interpretative nature** of social sciences and the role of logical fallacy in justification. Later on, it was extended by the introduction of the concepts of pragmatism, critical theory, and pluralist thinking (Bohman, 2002).

Tang (2011) argued that there were 11 foundational paradigms in the social sciences. He identified nine bedrock paradigms: (i) materialism, (ii) ideationalism, (iii) individualism, (iv) collectivism, (v) anti-socialization, (vi) socialization, (vii) biological evolution, (viii) harmony, and (ix) conflict, and two integrative paradigms such as (x) social system paradigm and (xi) the social evolution paradigm. Each of these can only shed light on a limited area of human society. Therefore, in order to understand human society and its history adequately, all of them should be deployed. Social sciences have always been multiperspective and multicultural in nature. This gave floor to various claims about how it is best to investigate and to understand the social world. Philosophy of social sciences deals with the generalized meaning of things and attempts to consider notions such as objectivism, normativity, replicability, quantifiability, explanation, demonstration, and prioritization.

Social sciences are centered on the **sharing of experience** about the social world in which perspectives of people differ from one another. Valsiner (2019) stated that social sciences are crucial in our understanding of the increasingly globalizing ways of living in the twenty-first century, which is characterized by the conflict of rapid technological advancements and the resistance of the traditional social orders to them. Reber and Bullot (2019) discussed the difficulty of drawing a clear demarcation line between science and evidence-based advocacy in the social sciences and humanities. They identified several open research issues such as (i) motivated testing, (ii) including and weighting values, (iii) side effects, (iv) intuitive judgments, (v) relativism and reductionism, and (vi) conditional objectivism. For instance, conditional objectivism claims that researchers have to (i) acknowledge value plurality, (ii) consider multiple standpoints for drawing practical conclusions, and (iii) reason based on counterfactual conditional statements.

As philosophical positions with remarkable influence on community psychology, Tebes (2017) considered: (i) perspectivism, (ii) pragmatism, (iii) feminism, and (iv) critical theory. These positions (i) seek to base their claims on empirical evidence, which is accepted within a given scholarly community, (ii) accept constructivism as the basis for knowledge claims to varying degrees, (iii) recognize that knowledge claims are flawed and dependent on culture, history, and unique contexts, and (iv) seek to use knowledge claims variously as the basis for action. Gamma-disciplines have things in common with sciences as well with humanities. The goal of social science is to understand and explain social phenomena around us, of which we ourselves are a part. Compared with the beta-disciplines, social sciences seek knowledge differently and offer different knowledge. For example, instrumentalists commit themselves to the view that social sciences, like engineering, should conduct only applied research and should devote its capacities to the creation of innovative solutions for real-world problems.

Certain fields of social science are often criticized as unscientific because of their limitations in formulating general laws and universal theories governing human societies. Mayntz (1990) posited that natural science concepts and models have had significant influence on theoretical developments in the social sciences, especially on sociology, in terms of building of theories, formal modeling, and transfer of theories. Driven by the intention of establishing a common physics foundation for all fields of social and natural science, Wayne (2013) made an attempt to formulate five new physics laws, which are qualitative, related to decision-making, and connect particular research domains of natural and social sciences. Rittel and Webber (1973) argued that a significant difference between natural science and social science is that the former has developed to deal with "tame" problems, whereas the latter usually faces "wicked" problems that cannot be definitively described.

4.5. Delta knowledge

Some elements and levels of constructivism have appeared in gamma-disciplines and -knowledge. On the other hand, the operationalization and utilization of **truly creative knowledge** is rooted in intellectual aptitudes such as heuristics, intuition, instinct, serendipity, inception, trial-error, perception, clairvoyance, and karma. These words implicitly express interrelatedness of the potency of creating and the intention of creating that are closely related to humans and their designerly behavior (Gedenryd, 1998). They are not about what is, but about what might be. This explains why they have not been used in the literature in the context of ABG disciplines. Having recognized the unique characteristics of design knowledge, Gilles and Paquet (1989) suggested differentiating socio-cognitive inventive knowledge as a fourth genre. Referred to as delta-knowledge, this genre supplements the knowledge of the ABG disciplines. It includes the intellect of various creative disciplines such as fine arts, performing arts, industrial design, product design (customer durables), as well as that of sustainability, law, and economy (Giard and Gilles, 2001).

Design science has been defined as a body of knowledge produced by rigorous research about what, why, and how to design. What makes (engineering) design research different to that of science is (the necessity of) abandoning the explicit commitment to evidence, though not neglecting physical principles and critical thinking (Eekels and Roozenburg, 1991). In the process of design, evidence about the properness of the outcome may not be logically possible since the design problem and the designers' intellect evolves with the design solution, and vice versa (Frayling, 1993). Another essential characteristic of **design knowledg**e is that it is closely related to disciplinary practice. As stated by Friedman (2017), design practice is a significant method of creating new knowledge and deepening existing knowledge by practitioners.

Delta-knowledge differs from common-sense knowledge since it has rich professional content. It also differs from alpha-knowledge because of its dynamically evolving nature. With regard to beta-knowledge, the main difference lies in the different roles of and relationships to formal theories, and in the dominance of validation in context over logical justification of proper beliefs. In comparison with gamma-knowledge, the difference is in the dominantly creative and predictive nature, and not in a descriptive and explanatory nature (Van Aken, 2005). Delta-knowledge shares the **pragmatic nature** of gamma-knowledge, and reflects the particular, rather than the general. It may exist in two forms. First, it can manifest in a noticeable higher sophistication (or quality) of the subject of making, doing, acting, and deciding, in particular, when the mentioned activities occur in a recurring manner. Second, it can also manifest in the mind-set and skills of the maker, doer, actor, and decision-maker and may enable a successful task completion, but it conflicts with beta-knowledge neither on a fundamental nor on an applied level. It is gaining an important place as part of the transdisciplinary innovation assets of change-maker disciplines (Cross, 2001).

Systematic inquiries toward design knowledge may address any phenomenon related to (i) artifacts (products, systems, services, and experiences) created by design, (ii) people (involved in or influenced by design), (iii) processes (involving all creative, operational, use, and change activities), (iv) environments (in which design-related changes take place), and (v) intellect (cognition associated with intelligent behavior). In each of these categories, knowledge may ideate, describe, explain, predict, and/or regulate natural or created phenomena. In addition to dedicated inquiries, a basic mechanism of acquiring delta-knowledge is generalization from practical cases and situations, as direct or indirect **reflection in action** (Schön, 1988). Direct reflection assumes a conversation with the subject and the situation, and an intuitive or systematic evaluation of the findings and experiences. Indirect reflection focuses on the implications of design decisions and dealing with wicked-problems that often happen in design, policy, and planning (Graham and Dickinson, 2007). From a structural perspective, delta-knowledge is composite knowledge. It purposefully blends knowledge of (i) natural sciences, (ii) engineering sciences and technology, (iii) social and behavioral sciences, (iv) creative and applied arts, (v) humanities and liberal arts, and (vi) human professions and services.

Other issues of delta-knowledge are related to understanding its typology and possible taxonomies. For instance, Uluoğlu (2000) approached the type of design knowledge from the perspective of communication and distinguished (i) reflective, (ii) operative, (iii) contemplative, (iv) directive, and (v) associative types of knowledge. As fundamental ones, Narváez (2000) considered (i) empirical– analytical, (ii) hermeneutical–historical, and (iii) socio-critical types of knowledge. Cross (2001) distinguished three categories of design knowledge: (i) knowledge of people (of outstanding designers), (ii) knowledge of design processes, and (iii) knowledge of design artifacts. Van Aken (2005) made a distinction between (i) object knowledge (knowledge about properties of the artifacts and technologies), (ii) ideation process knowledge (knowledge about the design processes to produce object or realization designs), and (iii) realization knowledge (knowledge about the processes to realize artifacts). Considering system cognition, Radermacher (1996) proposed a general knowledge framework that includes four levels: (i) physical, (ii) neuronal, (iii) symbolic, and (iv) model knowledge levels.

4.6. Epsilon-knowledge

The primary goals of the fifth industrial revolution are to develop artificial narrow and general intelligence as a **society-level productive asset** and to utilize it in productive systems. Based on what the current trends project forward, many experts expect disruptive changes in the twenty-first century (Sheng et al., 2019). Partly, such changes will be caused by i*CPSs and other resembling systems. As mentioned earlier, a strong driver of this is the on-going and intensifying trend of blending fundamental entities and processes (Canton, 2004). As the BANGM technologies evolve, a true synthesis of the physical, biological, cognitive, digital, cyber, and social realms becomes possible. Through this synthesis, various levels of intelligence also become naturally integrated into engineered systems. This creates opportunities for basically different systems and applications beyond artificial narrow intelligence systems, collaborative agent-based software systems, and smart cyber-physical–social systems. As a result, intellectualized (a.k.a. smart) and highly-intellectualized (a.k.a. intelligent) engineered systems will be available for numerous conventional and unconventional applications (Liu et al., 2004).

What makes highly-intellectualized systems different to well-known knowledge-intensive systems is their continuously evolving (not fully pre-programmed) system intellect (intelligence). The intellect of systems evolves through the growing amount of synthetic knowledge acquired by them and the sophistication of the related ampliative reasoning mechanisms. Systems equipped with these resources will be able to aggregate, produce, learn, transform, employ, and experience intellect over time, in addition to using their initially existing intellect in **system-level problem-solving** (Sumari and Ahmad, 2017). Considering these facts, the major claim has been that the massive SSK, which is generated and aggregated by multiple systems (such as intellectualized cyber-physical system of systems), is going to grow into a fifth genre of knowledge. It will be a fully-fledged complement of the ABGD-genres of knowledge. Therefore, it can be called "epsilon-knowledge" (Figure 5.2). The rapidly improving knowledge generation and acquisition capabilities of intellectualized engineered systems intensively support dynamic formation of this **new genre of knowledge**.

Epsilon-knowledge includes the total of knowledge associated with the operation of intellectualized engineered systems either as system-level problem-solving or systems state maintaining knowledge. It has its own features, methods, and appearances. The main constituents of massive SSK are (i) codified human knowledge (pre-programmed in individual systems), (ii) illative/inferential knowledge (self-generated by individual systems), and (iii) aggregated meta-knowledge (generated based on the contributions of the linked individual systems). Codification of human knowledge establishes an implicit (interpretative) relationship between genuine human beliefs and system knowledge. The process of codification includes aggregation, filtering, structuring, representation, and validation of the raw knowledge elicited from relevant human stakeholders. In combination, it involves both external and internal knowledge engineering. The latter has dominance after the booting up stage of systems. This is a typical situation in the case of using digital twins related to i*CPSs (Boschert et al., 2018). If the internal knowledge engineering process is an intense one, then a rapid growth of SSK can obviously be anticipated. This means that the actual system knowledge will be very different from the start-up knowledge after a longer period of operation. A practical example is automated training of deep

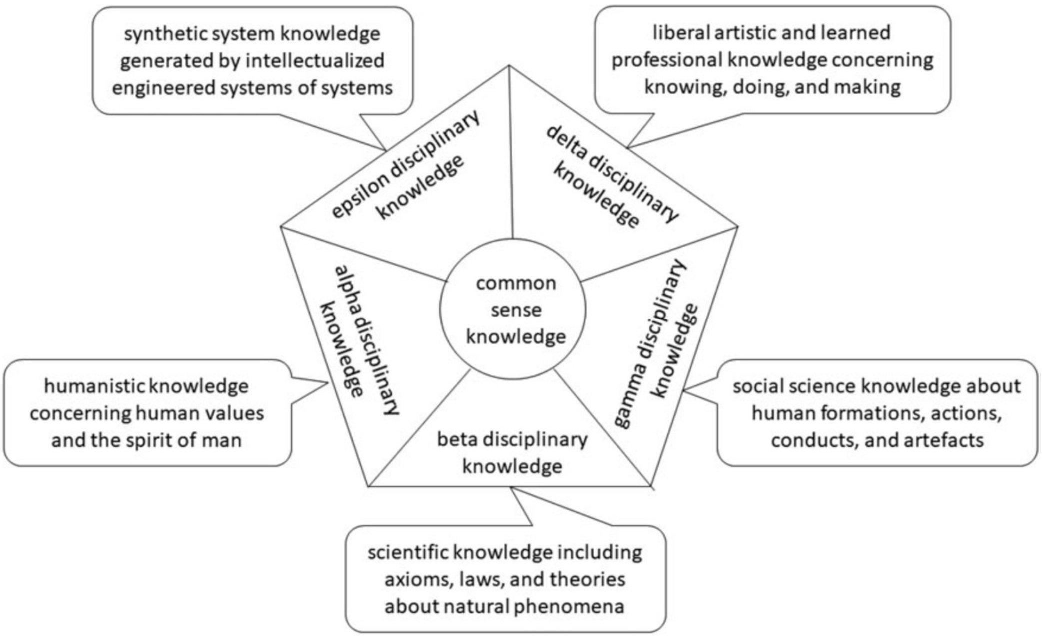

Figure 5.2: Epsilon-knowledge as a fully-fledged complement of the ABGD-genres of knowledge

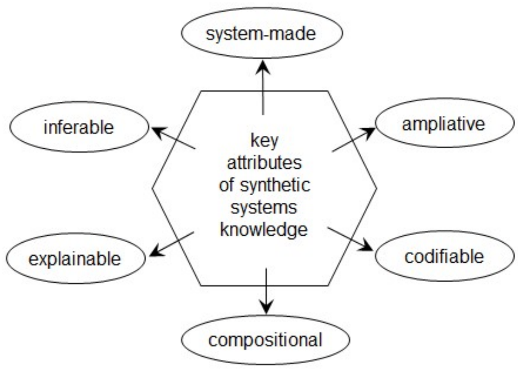

Figure 5.3: Key attributes of epsilon-knowledge

learning neural network mechanisms.

In addition to generic attributes (such as transferable, traceable, duplicable, augmentable, evolvable, and experiencible), SSK has **distinguishing key attributes**, namely: (i) system-produced, (ii) codifiable, (iii) ampliative, (iv) compositional, (v) explainable, and (vi) inferable. These unique and exclusive characteristics are shown in Figure 5.3. The term 'system-produced' means that this knowledge is essentially artificially (synthetically) created by systems for given purposes by complementary knowledge generating and fusing actions. It can be explicit (like a rule base) or implicit (like a model learnt by machine learning). The original start-up information (like training data) has a minor or no role in the whole of SSK. 'Codifiability' represents the opportunity and extent to which a given knowledge item can be reduced to information by means of drawings, formulae, numbers or words. The attribute 'ampliative' describes the ability to derive additional knowledge that is explicitly not included in a given body of SSK. The attribute 'compositional' implies that the constituents of SSK are functionally and/or semantically (cognitively) dependent on each other and are thus inseparable. The attribute 'explainable' refers to the possibility of discovering, analyzing, and clarifying explicit or implicit nonfigurative relationships between unconnected, causal, and/or abstracted constituents of SSK directly or indirectly. Lastly, 'inferable' means that SSK complies with the principles of epistemic knowledge (logical rationality) and allows multiple forms of reasoning without the need for a system to know explicitly what it knows implicitly.

The chunks and bodies of SSK may be represented explicitly and implicitly. Explicit knowledge representations capture the functional elements and the logical and semantic arrangement of knowledge by means of language constructs and procedural structures. Implicitly knowledge representations are abstract information patterns or models generated by computational learning and reasoning mechanisms (Ziori, 2004). The forms of representation are inseparable from the kinds of the associated **computational mechanisms**, which are available for application problem solving and/or for self-construction and management of synthetic knowledge. The latter needs dedicated knowledge, which complements the problem-solving knowledge. In addition to intra-system mechanisms, also inter-system knowledge aggregation and processing mechanisms can also be expected in the near future. The cognitive mechanisms for self-management include computational mechanisms for context modeling, awareness building, situation analysis, decision-making, and communication management, which are naturally given in the case of human cognition.

4.7. Computational construction of epsilon-knowledge

The traditional way of constructing system knowledge is pre-programming filtered and structured human knowledge by a team of knowledge engineers. The approaches of preprocessing and coding human knowledge are commonly known from the knowledge engineering practice of knowledge-intensive systems (Aikins, 1983). Beyond these, several ways of producing SSK from a variety of sources have been developed. They can be sorted into two categories: (i) acquiring, inferring, learning, and managing knowledge by system-specific computational mechanisms, and (iii) aggregating and deriving meta-knowledge by cross-systems computational mechanisms. As far as the sources used for synthetic knowledge generation in i*CPSs are concerned, actually anything can be a source presumed it can be accessed and processed by some computational mechanisms. Figure 5.4 shows the most important **general content sources**.

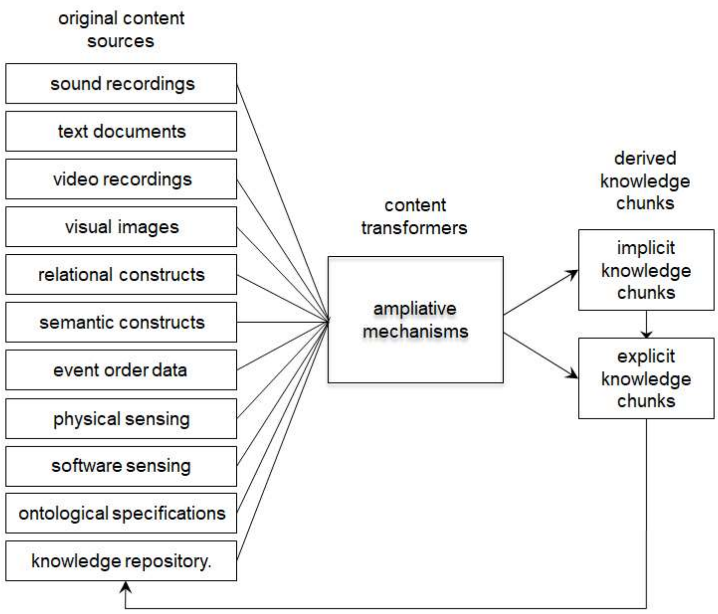

Figure 5.4: Examples of sources of data, information and knowledge contents for i*CPSs

The ampliative mechanisms of the system transform the contents (signals, data, and relations) elicited from the sources into problem/task specific or generic knowledge. This content-transformation process results in new knowledge chunks, which appear either in explicit or implicit forms and are deposited in knowledge repositories. Thus, the derived knowledge chunks can be used immediately as input for reasoning. The direct coupling between the sources and the repositories makes the content transformation process recurring. Well recognized challenges of a fully automated knowledge elicitation and processing are (i) diversity and variability of the representations of the contents obtained from the mentioned sources, (ii) extracting the semantic meaning in a context-sensitive manner, (iii) efficient transformation of the input contents into useful chunks of knowledge, (iv) synthesizing the chunks into coherent bodies of knowledge, (v) assigning meta-knowledge to the various bodies of knowledge, and (vi) the need for interoperable computational mechanisms, which may or may be not linked procedurally and semantically.

The mentioned general knowledge sources offer a wide range of input for i*CPSs development as well as for knowledge transformation. For example, analog and digital sound recordings include noise, speech, and music signals and data. Analog and digital text documents may include traditional text carriers and local digital or network hypertext files. Having both visual and auditory contents, video recordings may be real-time streams or stored recordings. Visual images include drawings, images, photos, and displayed contents. Relational constructs are entries of digital databases and digital models, while semantic constructs are arrangements such as scripts, frames, decision tables, agent intellect, and rule structures. Event order data represent logical and temporal (historical) relationships among events. Physical sensing provides descriptive characteristics (signals and data) of physical phenomena, while software sensing provides data associated with computational phenomena and actions. Ontological specifications carry descriptive and associative characteristics in structured (or standardized) language formats.

The cross-systems computational mechanisms of complex system-of-systems (SoSs) will extract, structure, consolidate, and store meta-knowledge in online warehouses. Toward this end, their operation may resemble what is done by **collaborative multi-agent systems**. As simple examples of meta-knowledge orientated knowledge processing, we may consider: (i) a smart parking assistance system, (ii) indoor fire evacuation system, or (iii) building-integrated vertical greenhouse system. In the case of the smart parking assistance system, the problem-solving computational mechanisms select, optimize, and apply the parking strategy, whereas the meta-knowledge deriving mechanisms can learn the features of the successfully chosen and applied motion paths, or can rank their appropriateness with regard to concrete cases. Based on this knowledge, problem-solving mechanisms can be improved to operate more efficiently in future cases. In the case of a smart in-house fire evacuation system, the problem-solving computational mechanisms can build situation awareness, develop an effective evacuation strategy, and work out a smartphone-based informing plan, while the meta-knowledge deriving mechanisms can learn the rate and way of obedience of the people to the obtained information/instructions and offer this information for optimizing the messages sent to them real time and the whole messaging the during the evacuation process. In the case of the building-integrated vertical greenhouse system, the problem-solving mechanisms can learn the patterns of the variable sunlight, lighting, and humidity conditions and adjust the irrigation accordingly, whereas the meta-knowledge deriving mechanisms may monitor the consumption habit and optimize the planting and growing accordingly, or take overuse or underuse into consideration in planning.

4.8. Reflections and open issues

Jashapara (2010) stated that knowledge is "an intrinsically ambiguous and equivocal term". This chapter argues that SSK is not covered by the four genres of knowledge which have been known so far. The conducted literature study underpins this claim. It is proposed to regard this rapidly increasing knowledge as a new genre, called epsilon-knowledge. The dissertation refers to sympérasmology as the proper conceptual framework for studying this genre of knowledge. In simple words, what is argued in this chapter is that SSK must be regarded as an emerging complement of Machlup's types of disciplinary knowledge. This knowledge shows a specific development pattern. At the beginning of the aggregation process, this knowledge is "triggered" by some pre-processed (filtered, structured, represented, and coded) chunks of human knowledge. During the operation of intellectualized systems, this knowledge is supposed to be significantly extended, even possibly replaced by the associated computational mechanisms (Epstein et al., 2018). Based on the research, the key propositions are as follows:

- It is important to differentiate epsilon-knowledge from the knowledge that is used by the developers in designing and architecting the application specific reasoning mechanisms and knowledge repositories these systems. This latter is beta-/delta-knowledge, rather than epsilon-knowledge. An example of this delta-knowledge is the data constructs describing the feature sets used for training a deep-learning neural network.
- Epsilon-knowledge has different features than the bodies of knowledge that are related to human inquiries in the alpha-, beta-, and gamma-disciplines as well as to those related to human inventions in the delta-disciplines. It is associated with the evolution of intellectualized engineered systems that are made capable of reasoning with digitally coded human knowledge and to acquire, synthesize, learn, extract, aggregate, and restructure knowledge on their own.
- Epsilon-knowledge is influenced by human knowledge, decisions, and designs only implicitly and in a restricted manner. Its uniqueness originates in that it is produced for application problem solving and optimization of system performance.
- Epsilon-knowledge is differentiated by six key attributes: (i) inferable, (ii) system produced, (iii) ampliative, (iv) codifiable, (v) compositional, and (vi) explainable. They can be projected onto all forms of SSK.
- Epsilon-knowledge is sufficiently explored and explained yet neither from an ontological

nor a methodological point of view. Therefore, it needs further systematic investigations. On the other hand, it may become an additional productive asset if sufficiently understood and supported by management and exploitation strategies.

- Intellectualized engineered systems should be seen in the near future not only as AI-enabled problem-solving systems, but also as knowledge growing and harvesting systems. This will increase their functional complexity, but it will also offer new affordances and business advantages.
- Managing epsilon-knowledge implies the need for semantic and pragmatic knowledge fusion frameworks and system-independent methods and meta-methods, well beyond the issues of unification and conversion of knowledge models and representations.

Chapter 5

5. An approach to investigation of system-level problem-solving knowledge

5.1. Research objectives and approach

Explanation, typifying, and examination of the essential features of the types of human knowledge has a long historical tradition. There are abundant definitions and interpretations of human knowledge that has been categorized according to a large number of aspects. One genre is called **common-sense knowledge** (or individual knowledge). The philosophical stance of this kind of knowledge was addressed extensively over the centuries, but it is still the subject of many ongoing discussions in the specialist literature. As a broad and multi-faceted genre of human knowledge, individual knowledge includes three basic categories: (i) empirical knowledge (obtained by sensing, sensations, signals, trial-errors, measurements, observations, etc.), (ii) intuitive knowledge (obtained by beliefs, conjectures, self-evident notions, gut feelings, faiths, acquaintances, conventions, etc.), and (iii) authority knowledge (obtained by social norms, expertise, codified constructs, de facto rules, propaganda, ideology, etc.). A shared characteristic of these categories is that they are closely connected to human individuals or groups of individuals. The mentioned categories of knowledge are also characterized by (i) incompleteness and uncertainty, (ii) strong subject-dependence, (iii) situation-sensitiveness, (iv) context-reliance, (v) time-relatedness, and (vi) limitations in generalization. They are not, or not completely, factual and teachable forms of knowledge.

Another genre of human knowledge is **generalized knowledge** (which is generic, complete, justified, and universally applicable in context). This knowledge is mainly related to scientific inquiry and is the subject of epistemic thinking, which is committed to the scientific study of perceptive, cognitive, and linguistic processes by which knowledge and understanding are obtained and shared. Rigorous criteria are set to distinguish scientific knowledge from common knowledge, sensations, memory, introspection, and reasoning, though these are regarded as the ultimate sources of belief. Opposing weakly grounded knowledge, scientific knowledge is a complex formation that includes laws of nature, empirical facts, tested theories, formal models, speculations, and hypotheses, all originated and formulated by humans. These all coexist, evolve, and compete with each other. That explains why gnoseological and epistemological efforts coexist and strive for a deeper and more complete understanding of the genres of human knowledge. Seminal publications agree that knowledge is a productive asset, no matter if human knowledge or system knowledge is concerned. The same viewpoint is taken further in this dissertation.

Owing to the results of research in system science, artificial intelligence, and cognitive engineering, engineered systems are becoming more and more powered by SSK. Both the possibilities of generating epsilon-knowledge and the amount of SSK captured by intellectualized engineering systems are growing. Complementing common-sense and scientific knowledge, SSK is maturing into a crucial productive asset. These conclusive claims were stated at the end of the preceding chapter. It was also mentioned that an overall theory of epsilon-knowledge does not exist yet. Theoretical frameworks, explanatory theories, and methodological approaches for investigation and utilization of this genre of knowledge are not offered by the literature. In turn, these deficiencies have an objectionable effect on the status of SSK. Nevertheless, forward-looking researchers have argued that time has come to establish a philosophically underpinned theoretical framework. Their motion is legitimated by the on-going intelligence revolution, in which artificial intelligence becomes a productive power, a primary enabler of smart systems, and a strong transformer of social life.

In line with the above motion, systematic investigations have been made concerning the needed theoretical framework and, even further, towards a possible **new branch of philosophical studies**. The expectations for this theoretical framework were to outline and support the study of the overall nature, specific characteristics, internal relationships, probable impacts, ways of exploitation, and the future role of epsilon-knowledge has also been conceptualized. These results and other findings are presented in this chapter. First, the logical studies of the types of human knowledge and existing notional platforms for rigorous scientific studies of the various types of human knowledge are reviewed. Light is cast on their limitations from the viewpoint of the specific characteristics of SSK. Then, the proposed novel reasoning framework is introduced, including its origins and assumptions. Its foundational concepts are clarified and the primary investigation domains are discussed. The last part of the chapter projects the aspects of investigation to SSK and offers additional information about it. The contents of this chapter have been compiled based on the following peer-reviewed publications: H12, H14, and H15 (see Appendix A.1.1).

5.2. Gnoseological study of individual human knowledge

Historically, the Greek word “gnosis” has been used as an idiomatic term to name the common forms of “knowing” (Nguyen, 2015). This term is said to have multiple meanings, such as: “individual knowledge”, “perceptual knowledge”, “acquaintance”, “explicit opinion, belief, and trust”, and the Greek distinction of “doxa”. From the words “gnōsis” (for knowledge) and “logos” (discussion), the term “gnoseology” was coined to name the various studies concerning what can be known about common things and practices in a truly concrete sense. In the 18th-century, the Latin term “gnoseologia” was used by Baumgarten (1986) to name the study of non-teachable and not objectively testable knowledge related to aesthetic values and positions. For some scholars, (classical) gnoseology is the metaphysical theory of knowledge or the metaphysics of truth and it became coextensive with the whole of metaphysics. Lately, gnoseology has also been defined as (i) the philosophy of knowledge, (ii) the philosophic theory of knowledge, (iii) the theory of human faculties for learning. In addition, it has been seen by many scholars as the theory of cognition, due to its lack of interest in foundations, generalizations, and abstractions. Gnoseology has been focused on socially premised and historically loaded human sensory and affective **cognition**, viewing it as a process of achieving knowledge, the highest form of which is science. Consequently, one can also understand gnoseology as a way of knowing practical knowledge, reflexive knowledge, local knowledge, etc. without general and absolute cogency and validity. As such, it could even cover understanding of knowledge gained though meditation.

There are four aspects of inquiry in which the philosophic theory of knowledge is interested, namely: (i) the basis of knowledge, (ii) the nature of knowledge, (iii) the validity of knowledge, and (iv) the limits of knowledge. Gnoseology goes beyond “hermeneutics”, that is, the interpretation of beliefs, but does not concentrate on “epistemics”, that is, on evidential justification of beliefs. At large, it focuses on the **socio-cognitive aspects** of common individual knowledge, and addresses (i) the process by which the subject is transferred to a state of knowledge, (ii) the human faculties for perceiving, thinking and learning, (iii) the universal relationship of common knowledge to reality, (iv) the conditions of its authenticity and truthfulness, (iv) the role and manifestation of human cognition, and (v) the preconditions and possibilities of cognition. Though it has a tight connection with the theory of cognition, it differs from it (i.e., from the study of the mental processes and information generation and processing of the mind). In the interpretation of Li (2013), a gnoseological study may deal with people's spiritual and psychological phenomena, such as impressions, emotions, and meanings, in particular with people's value psychology, value perception, value concepts, and value evaluation. Interestingly, Nikitchenko (2011) proposed to use a gnoseology-based approach for developing methodological, conceptual, and formal levels of foundations of informatics.

Though the history of gnoseology is relatively long, it has not become common in philosophy and education for reasons that are difficult to uncover. The term is almost never used by English language philosophers. For the reason that intuitive and/or instinctual knowledge plays a less significant role in the productive segment of the industrialized society, the original concept of gnoseology too plays a less significant role. There are publications that refer to gnoseology as the theory of **non-human-rooted knowledge**. As a conclusion, gnoseology is primarily concerned with non-universal (i.e., particular, incomplete, non-justified) knowledge. As it focuses on non-scientific knowledge (or the whole complement of scientific knowledge), it is often seen as the theory of partial approximate knowledge. It addresses scientifically not-justifiable domains of knowledge, such as instinctual, heuristic, intuitive, common-sense, instinctual, sensual, axiological, experienced, etc. knowledge of individuals. Besides concrete individual beliefs, gnoseology has many source fields, such as daily life, social interaction, culture, education, religion, politics, authority, and profession, from the perspective of what it means to "know" in these fields. My understanding is that gnoseology can be deemed to study knowledge from a priori point of view, and epistemology from a posteriori point of view. In this vein, the concept of gnoseology is a broader and looser concept than that of epistemology.

5.3. Epistemological study of general human knowledge

Epistemological thinking may transcend gnoseology by means of its level of analysis, as it takes care of the specific conditions of production and validation of scientific knowledge (historical, psychological, sociological circumstances, and justification criteria). The exact criteria to distinguish generalized (scientific) knowledge and common (individual) knowledge are not really known (where does qualitative knowledge become scientific?) and the border line is blurred. Historically, we cannot find traces of modern epistemology until the last two centuries. This is in line with the consolidation of science as a social establishment and clarifies why epistemology has become a more operational concept over the centuries.

Introduced by the Scottish philosopher, Ferrier (1866), the term "epistemology" refers to the theory of well-grounded (testable and learnable) knowledge, which has been captured by the Greek word 'episteme'. In a narrow sense, epistemology is understood as the study of the conditions of production and validation of scientific knowledge. It accepts beliefs as the basis and necessary condition of a scientific system of knowledge, though certain epistemologists argue over whether belief is the proper truth-bearer. In our post-modern times, it is interpreted both as the **branch of philosophy** that studies the origin, nature, kinds, validity, and limits of human knowledge, and as the **theory of knowledge**, especially with regard to its scope, methods, and validity, and the distinction between justified belief and personal opinion. Epistemology is normative and critically evaluative, rather than descriptive and explanatory. Goldman (1986) conceived two interrelated fields of epistemology: (i) individual epistemology (related to cognitive sciences) and (ii) social epistemology (related to social sciences). Epistemology accepts four theories of knowledge, namely: (i) connectionist (associative), (ii) cognitive, (iii) constructivist, and (iv) behaviorist theories. Connectionist approaches focus on the presence or absence of associations and their quantity, while constructivist theories contemplate the reasons of knowledge such as causality, probability, and context. Behaviorist theories interpret knowledge as behavioral responses to different external stimuli, but do not contemplate internal cognitive (thought) processes. In several cognitive theories, knowledge is treated as integrated and abstracted structures of information of various kinds. Hjørland (2009) proposed that there are four basic epistemological approaches to knowledge organization, namely: (i) empiricist, (ii) rationalist, (iii) historicist, and (iv) pragmatist.

As the theory of knowledge, epistemology has an extremely **wide range of concerns**. It poses questions and tries to find defendable answers to them. One family of questions is about the very essence (nature) of knowledge: (i) What is sensory experience? (ii) What are beliefs? (iii) What is the rationality of beliefs? (iv) How do beliefs turn into knowledge? (v) What are the criteria for knowledge? and (vi) What is the nature of knowledge? A second family is about how humans can come to know: (i) What are the sources of knowledge? (ii) What is consciousness? (iii) What is the role of memory? (iv) What is the truth of reason? (v) What are the kinds of testimony? (vi) How do we know that we know? (vii) What is the relationship between the acquired knowledge and the knower? (viii) What is the role of skepticism? and (ix) How does language construct knowledge?

A third family is about the conditions of knowledge: (i) What is justification? (ii) When is a belief justified? (iii) What makes justified beliefs justified? (iv) What is truth? (v) When is knowledge true? (vi) What is correctly proven true belief? What is correct inference and reasoning? (vii) When is knowledge coherent? (viii) What is dogmatism? (ix) What is the grounding of scientific knowledge? (x) What is moral knowledge? and (xi) What is religious knowledge? Lastly, a fourth family of questions is about the insinuations of knowledge: (i) When do we know something? (ii) What is fallibility? (iii) How do we ensure valid knowledge? (iv) What is belief-less knowledge? (v) What is common sense? (vi) What is evidence? (vii) What is abstraction? (viii) What is context? (ix) What are possibilities for knowledge integration? (x) What is intelligence? and (xi) What is wisdom? It is difficult to make any prioritization in terms of the questions as most of them are interrelated.

In the last two decades, epistemology-oriented thinking and striving for an **epistemology of complex systems** have penetrated into systems science and system engineering. Only some illustrative contributions can be mentioned here. Hooker (2011) argued that, ultimately, the goal of science philosophy is to develop mature foundations/philosophy of complex systems, but attempting this is premature at this time. Helmer and Rescher (1959) discussed the need for a new epistemological approach to the inexact sciences, since explanation and prediction in the case of these sciences do not have the same logical structure as in the exact sciences. As new methodological approaches, they mention system simulation and expert judgment. Ratcliff (2013) exposed a specific application of epistemology to support system engineering. Figueiredo (2008) argued that the developing epistemologies of design and engineering contribute to a renewed epistemology of science. Boulding (1987) concluded that the primary obstacle to the development of robust knowledge platforms for engineering systems is the lack of developmental and operational predictability. Möbus (1996) posited that **hypothesis testing** plays a fundamental role in a cognitive-science-orientated theory of knowledge acquisition as well as in handling problem-solving knowledge in intelligent problem-solving environments.

Based on the above considerations, the following can be concluded. As Figure 5.1 shows, gnoseology deals with (everyday, non-generalized) human individual knowledge possessed by individuals. Epistemology investigates generalized human knowledge, which manifests in the form of justified scientific knowledge. Epistemology may address

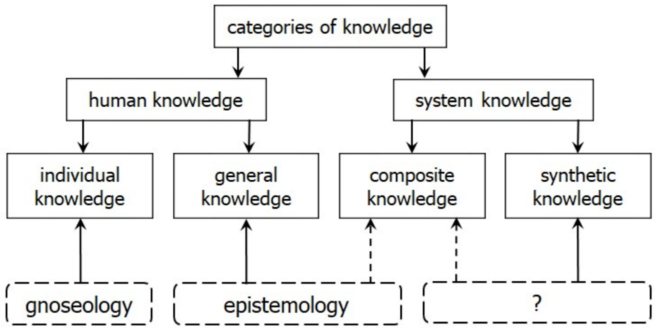

Figure 5.1: Studies of categories of knowledge

composite engineering knowledge, assuming that it is a combination of everyday and generalized human knowledge. However, neither gnoseology nor epistemology has shown interest in a fully-fledged study of dynamically changing SSK. Actually, they are not ready to host the emerging genre of epsilon-knowledge for teleological and historical reasons. The fact of the matter is that we are in a situation where the state of the art of systems knowledge is ahead of the state of understanding. Stokes (1997) argued, "knowledge finds its purpose in action and action finds its reason in knowledge". Thus, (synthetic) systems knowledge should be interpreted concurrently as an acquired cognitive capacity and as an actionable potential in context.

5.4. Revisiting the enablers of system intellect

Like human intelligence, system intelligence is a complex, multi-faceted, and yet not completely understood phenomenon and concept. Seeing it from outside, it stretches into two dimensions, as shown in Figure 5.2. The vertical dimension includes the two enablers of system intelligence (a dynamic body of application/operational knowledge and a set of associated reasoning mechanisms that purposefully alter and interact with each other while operating on this knowledge). The horizontal dimension includes the two sources of system intelligence (human-provided part and self-acquired part). Looking at them from inside, the reasoning mechanisms and the problem-solving knowledge are functionally and methodologically interconnected and thus inseparable. Knowledge and mechanisms are also computationally interconnected and make, for instance, CPSs capable of solving application problems and maintaining the efficiency of their operation. The synthetic intelligence primarily comprises self-acquired and/or self-generated application-specific knowledge and processing (reasoning) mechanisms, but it is not absolutely independent from the human-provided part. As mentioned above, I focus on synthetic system intellect in this dissertation, which is regarded as a subset of synthetic system intelligence.

Mentioned earlier, the knowledge possessed by systems typically blends (i) structured, formalized, and pre-programmed human knowledge, and (ii) knowledge that is acquired by a system during run-time using its own resources. The relative amount and significance of the human created initial (inputted) part usually decreases during the operation of systems. On the other hand, the relative volume and significance of the system-produced part - the authentic **synthetic system knowledge** (SSK) - grows through the useful life-cycle of systems. Thannhuber (2005) proposed to consider system knowledge both from a microscopic and a macroscopic perspective, which can be extended to the associated reasoning mechanisms too. Microscopically, knowledge is given by implementation level procedures or actable coordination processes (microscopic actions of a system). Macroscopically, knowledge is given by the

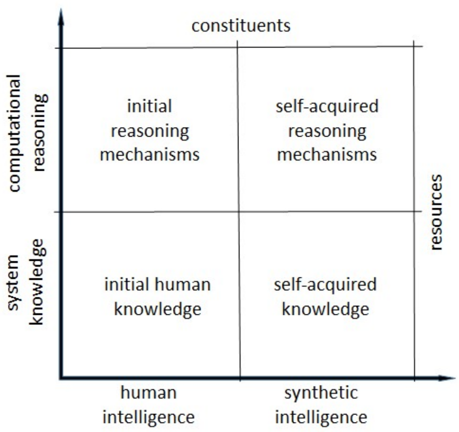

Figure 5.2: Dimensions of system intelligence

constraints and control of the declarative assembly mechanism that provides a meaningful system response to a given stimulus.

Synthetic system knowledge is not the knowledge that is needed to specify, design, implement, use, and recycle systems. Instead, it is the knowledge that is needed by systems to achieve the operational purposes or objectives, or in other words, it is the knowledge needed to function (solving problems). Thus, system knowledge is identical neither with engineering knowledge (the knowledge of making), nor with technological knowledge (the knowledge of enabling), though some elements of both are present in systems knowledge. System knowledge is a representative form of epsilon-knowledge including (i) chunks and bodies of generic scientific knowledge (facts, definitions, and theories), (ii) specialized professional knowledge (principles, heuristics, and experiences), and (iii) everyday common knowledge (rule of thumbs). A recently recognized issue is the balance between application neutrality/specificity and performance dependability/efficiency that concerns both the system knowledge and the processing mechanisms (Aussenac-Gilles et al., 2020).

Human reasoning is the progenitor model of computationally implemented inferring and reasoning processes. It is an intricate mental process of making logical conclusions and predictions from available knowledge in various application contexts (Stenning and Van Lambalgen, 2012). It can be both intuitive (heuristic) and formal (systematic), but both forms are influenced by the actor, purpose, problem-specific knowledge, and context information. The overall **formal mechanisms of reasoning** are underpinned by computational thinking. It is characterized by a logical procedure that involves the following steps: (i) specification (choosing and formulating a problem), (ii) decomposition (breaking a complex problem down to smaller and manageable sub-problems), (iii) patterning (identifying and representing a structure or a trend within the problem), (iv) abstraction (identifying specific similarities and differences among resembling problems), (v) algorithmizing (developing step-by-step instructions for solving the problem at hand), and (vi) analysis (reflecting on the characteristics of finding a solution for the problem).

With regard to computational reasoning, three core features are to be considered, namely, (i) moving from multiple inputs to a single output, which can be a conclusion or an action, (ii) making multiple steps through a state space to achieve a final outcome (in numerous ways), and (iii) processing the objectives, a mixture of previous knowledge, novel information, and the dynamic contexts (Mohaghegh and McCauley, 2016). Usually, **search space-based** (retrieval) and **additional content deriving** (ampliative) computational reasoning mechanisms (CRM) are differentiated. Ampliative CRMs are mechanisms that produce additional knowledge based on the knowledge externally embedded in or internally acquired by the system. Systems engineering also distinguishes system-level and constituent-level reasoning mechanisms.

The processing (reasoning) mechanisms can either be task-independent mechanisms or task-dependent mechanisms. Historically, five major families of computational mechanisms have been developed. First, symbolist approaches, such as (i) imperative programming language-based (procedures-based) reasoning, (ii) declarative logical language-based reasoning, (iii) propositional logic-based inferring, (iv) production rule-based inferring, (v) decision table/tree-based inferring. Second, analogist approaches, such as (i) process-based reasoning, (ii) qualitative physics-based reasoning, (iii) case-based reasoning, (iv) analogical (natural analogy-based) reasoning, (v) temporal (time-based) reasoning, (vi) pattern-based reasoning, and (vii) similarity-based reasoning. Third, probabilistic approaches, such as: (i) Bayesians reasoning, (ii) fuzzy reasoning, (iii) non-monotonic logic, and (iv) heuristic reasoning. Fourth, evolutionist approaches, such as: (i) genetic algorithms, (ii) bio-mimicry techniques, and (iii) self-adaptation-based techniques. Fifth, connectionist approaches, such as: (i) semantic network-based, (ii) swallow-learning neural networks, (iii) smart multi-agent networks, (iv) deep-learning neural networks, (v) convolutional neural networks, and (iv) extreme neural networks. Logical, retrospective, probabilistic, deductive, and inductive approaches represent the traditional approaches to the generation of deterministic knowledge. They have recently been extended by various learning and abductive approaches. Six main methods are known for reasoning with imperfect knowledge: (i) Bayes theory-based, (ii) Dempster-Shafer theory-based, (iii) fuzzy set theory-based, (iv) measure of (dis)belief theory-based, (v) inductive probabilities-based, and (vi) non-monotonic reasoning-based methods. Their application and operationalization in problem solving depends on the type and representation of the concerned chunks of the system knowledge.

System related knowledge may manifest in different forms and can be classified in multiple ways. The literature interprets: (i) declarative knowledge (which captures descriptors and attributes of facts, concepts, and objects), (ii) structural knowledge (which establishes semantic relations between facts, concepts, and objects, (iii) procedural knowledge (which captures the know-how of doing or making something), (iv) abstracted knowledge (which generalizes both “what is” and “how to” types of knowledge within or over contexts, (v) heuristic knowledge (which represents intuitive, emergent, uncertain, and/or incomplete knowledge in a context, and (vi) meta knowledge (which apprehends wisdom and decisional knowledge about other types of knowledge) can be differentiated as the main categories. The types of knowledge used in problem-solving are (i) declarative, (ii) procedural, (iii) schematic, (iv) strategic, (v) situational, (vi) metacognitive, and (vii) problem-translating knowledge. From the perspective of computational problem solving, (i) problem-environment (context) knowledge, (ii) problem-scheme (framework) knowledge, and (iii) problem-solving (content) knowledge have been identified (Solaz-Portolés and Sanjosé, 2008). Each one of these may be implanted into systems by knowledge engineers, or may be acquired and generated by the systems themselves during run-time. Proper syntactic or semantic **representation schemes** are needed to be associated with the individual knowledge categories. It is still a concern how representations can guarantee the synergy and compositionality of system-level knowledge. The related problem can be traced back to the unavoidable decomposition into primitives for computational processing.

5.5. Assumptions and objectives of sympérasmology

In order to initiate a specific theory of SSK, I proposed ‘**sympérasmology**’ as a comprehensive field of research and a conceptual framework for dealing with epsilon-knowledge. The name ‘sympérasmology’ has been derived by putting together the Greek term ‘**sympérasma**’ (which refers to inferred/reasoned knowledge), and the Greek term ‘**logos**’ (which is used to express the logic and reasoning in crafting a defendable piece of knowledge and the typically accompanying process of bringing up demonstrative logical cases). Besides referring to knowledge that is deductible and concluded, this term also has a second meaning, which expresses a bottom line of knowing or a consolidated conclusion. I took the liberty to use the term ‘intellectualization’ to refer to various grades of self-managed intellect in engineered systems. **Sympérasmological investigations** can help separate the overlapping concepts of system intelligence (i.e., intelligence reproduced by mimicking human intelligence) and system intellect (system intelligence by computational knowledge synthesis and reasoning) and the competing concepts of (self-managed) intellectualization and intelligentization.

The major assumption of sympérasmology is that an overwhelming part of the SSK is “sympérasma”, that is, knowledge conjectured, inferred, constructed, or otherwise derived during the operation of systems. This part is becoming dominant during longer operation of i*CPSs. Three basic requirements have been posed for sympérasmology. Namely, it must be: (i) intelligible, i.e., comprehendible by professionals working on SSK, (ii) robust, i.e., allow for the future developments of the domain, and (iii) distinguishing, i.e., should maintain its conceptual uniqueness and utility. The practical objective of sympérasmology is to become the staircase to the ‘mind’ (the inferred intellect) of IESs. Towards this end, it must address theoretical and conceptual issues of dealing with epsilon-knowledge in the context of system intelligence, rather than only methodological, technological, and engineering ones. It must consider logical, computational, semantic, pragmatic, apobetic, human, and social discourses.

Sympérasmology cannot be implemented as a simple augmentation of gnoseology or epistemology, since it is based on different fundamental concepts and principles originating in the very nature and processing mechanisms of SSK. The importance of sympérasmological studies of samples of epsilon-knowledge comes from the vision of knowledge growing and harvesting systems. Sympérasmology is supposed to provide a constructive and pluralistic theory of epsilon-knowledge. It must investigate its entirety, and not only the bodies of knowledge acquired by various system-level processing mechanisms. The issue is: if beliefs, truths, justification, foundationalism, and coherentism are the main conceptual pillars of contemporary epistemology, what will these be in the theory of SSK? In this sense, sympérasmological investigations go beyond the known methodological approaches (informed speculation, comparative analyses, and rational examinations) of gnoseology and epistemology. The methodological approach of sympérasmology may rely on methods such as: (i) critical literature studies, (ii) experimental investigations, (iii) prognostic systems thinking, (iv) cross-cases practical studies, and (v) contemplative validations.

The benefits of conducting studies according to the principles of sympérasmology are not only in a clearer academic view, but also in the opportunity of more dependable innovation strategies and better engineering decisions concerning the proper intellectualization of systems for industrial and social applications. It may even have a disruptive influence on the design, engineering, application, and utilization processes of smart (and intelligent) systems. It is foreseen that the sympérasmological theory of massive system knowledge will condense and evolve in accordance with the progression of systems science, systems technologies, and systems engineering. The sympérasmological insights will be **instrumental** for forwarding epsilon-knowledge towards a scientific, technological, and commercial asset.

5.6. Domains of sympérasmological investigations

Sympérasmological investigations may extend to a large number of interest-domains and may address many phenomena and problematics within and over multiple domains. A domain is interpreted as a specific topical area within a discipline. Domains of interest maintain a separation of the concerns and create a specific conceptual framework for organization of the inquiry efforts. The structural decomposition of sympérasmology into possible investigational domains is shown in Figure 5.3. These domains can be sorted into four categories called: (i) **rudiments**, (ii) **principles**, (iii) **faculties**, and (iv) **implications**. Due to the holism of SSK, many of the subjects/aspects of studies in the four domains are interrelated, dependent on the others, or even overlapping. The domain-specific phenomena and problematics are investigated from the perspective of holism, but attention is given to the dialectic relationships of convergences and divergences. Like in the case of epistemological studies, abundant sets of concrete inquiry questions can be formulated related to each domain.

The domain called ‘rudiments’ lends itself to the study of the basics, namely, (i) the nature and essence, (ii) the sources and channels, (iii) the models and representations, (iv)

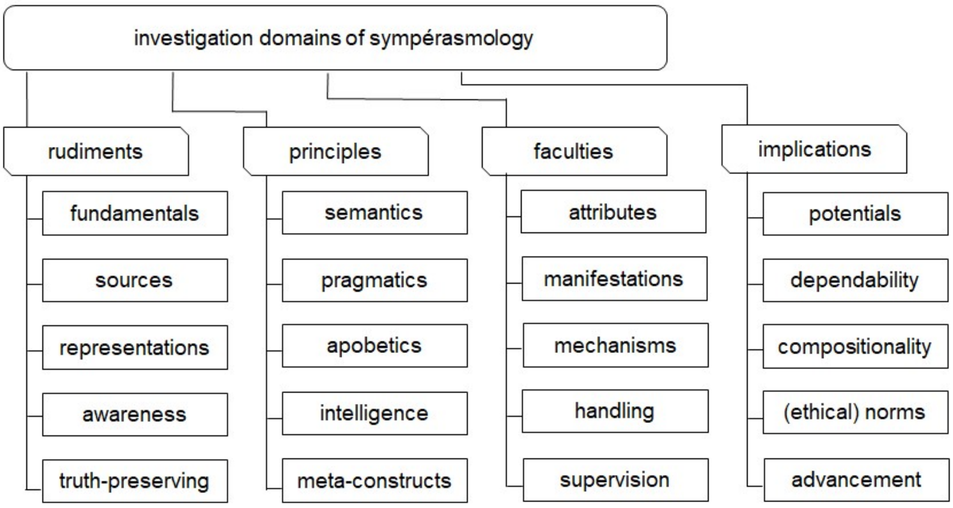

Figure 5.3: The proposed investigation domains of sympérasmology

the computational (non-human) awareness and cognizance, and (v) the properness in context of systems knowledge. The domain of 'principles' includes explorative and explanatory analyses into (i) computational interpretation and semantics, (ii) computational rationale and pragmatics, (iii) computational intentionality and apobetics, (iv) computational system intelligence (including aspects of learning, adaptation, and evolution), and (v) computational meta-constructs and notions. The domain of 'faculties' addresses SSK as an asset by focusing on (i) the distinguishing computational characteristics, (ii) the theoretical and practical manifestations in systems, (iii) the variety of processing mechanisms, (vi) the handling the synthesized knowledge in systems and across systems, and (v) the forms and limitations of human supervision.

The domain of 'implications' concerns inquiries about (i) problem solving affordances and potentials, (ii) reliability and dependability, (iii) compositionality (augmentation and integration), (iv) the ethical and other social norms, and (v) future progression and opportunities. The knowledge-centered investigations also consider the related cognitive processes and technological enablers. Sympérasmology is closely related to computing, but has a different viewpoint. As an example: the main issue for sympérasmological representation (syntax) studies is not what representations exist for different purposes and how to represent a particular body of knowledge digitally, but how the various representations relate to the very nature and essence of SSK. From a methodological viewpoint, sympérasmological investigations apply both 'a posterior' (experimental) and 'a priori'(interpretative) methods similar to those of epistemology and gnoseology. Normative questions are also posed concerning what and how people should view and approach SSK. The typical experimental method is individual case or case ensemble implied reasoning.

5.7. Rudiments of synthetic system knowledge

The word 'rudiments' is used to express the basics and essentials of the systems knowledge phenomenon, which make its existence, identification, treatment, and utilization possible, and from which other concepts are derived. Evidently, the rudiments are developing due to the changes in the SSK phenomenon and may be reformulated due to the unceasing progress of knowing. The essence of epsilon-knowledge (synthetic system knowledge) is empowering and endowing engineered systems with problem-solving power (rational intellect). Based on this, an intellectualized system can select the best sequence of actions leading to one of its goals as directly as possible. As handled by the system, this knowledge is rather a blend of problem-solving affordances, than a roadmap to or a receipt of arriving at a solution. Several publications demarcate (i) domain knowledge (aggregated factual knowledge about domain entities, attributes, relations, etc.), and (ii) problem solving knowledge (including procedural, reasoning, and regulatory knowledge) (Bratianu and Andriessen, 2008). With regard to sources of knowledge, an issue is that current systems do not have conscious beliefs in the same way as humans do. The sensors, memories, and/or algorithms are not true equivalents of the sources of natural human knowledge. They are primarily technological means, which do not stand in the same synergetic relationship with the processors/software, as the human senses are with the mind, and processors cannot be compared to the brain. This means, different than analogical contemplation is needed to identify and explain the sources of SSK.

Ultimately, systems knowledge is captured in formal, in-process representation using digitally manageable constructs. This representation is not working knowledge, though it is supposed to convey both the proxy meaning and the problem-solving potential of it. While syntactic representations of data and information has reached a level of sophistication, knowledge level representations are still suffering from the implications of the necessity of syntactic formalization and the limitations of capturing the intrinsic working potential of problem-solving knowledge in a direct manner. A context-dependently actionable problem-solving potential requires purposeful inferring from computable representations. Main issues for sympérasmological studies are how: (i) representations relate to the nature and essence of systems knowledge, (ii) syntactic representations can be exceeded, (iii) descriptive, relational, procedural, abstracted, heuristic, and meta-knowledge representations can be unified, and (iv) the problem-solving potential can be optimized.

As knowledge is a self-conscious act, consciousness, self-awareness, and a sense of self are supposed to be explained and approximated in intellectualized engineering systems. However, this cannot be studied by way of analogies. As for now, reproduction of natural human consciousness is deemed to be a mission impossible for reasons such as (i) there is no real demand, only weak curiosity behind it, (ii) it cannot be studied as a phenomenon by traditional reductionist approaches of empirical science, (iii) no transdisciplinary theory is in formation that would be able to explain all aspects of its formation/emergence beyond all question, (iv) its computational implementation would probably need resources that not available today. That is why the issue of non-conscious awareness is an important one for sympérasmological studies. Awareness is not only having information about something, but also being perceptually, cognitively, and emotionally involved in that. It is the ability of systems to make their behavior proper and rightly dependent on some knowledge and the environment. A challenge in this context is that awareness and sentience (the capacity of feeling, perceiving, or experiencing subjectively) are in a close coupling, but computational sentience is almost neglected in system\ studies.

Trueness and preserving truth in reasoning in changing context is currently the subject of both pragmatic and philosophical debates. It is unclear what makes SSK 'justified' or 'warranted' and how this can be achieved by systems themselves. If needed, what can replace the causally grounded, justificationally grounded, and epistemically grounded principles of justification in intellectualized systems? Pragmatists claim that truth of systems knowledge forms a relative category, which can only be vindicated in terms of its properness for the purpose it serves. However, it concerns both intentional properness (due to the changes in purposes/problems) and conditional properness (due to the life cycle and obsolesce of knowledge).

5.8. Principles of synthetic system knowledge

The term 'principle' is used to express the assumed relationships among the inherent characteristics and the observable manifestations and performances of SSK. Principles serve as the foundation for reasoning. For instance, forming meaning and making sense are essential characteristics of SSK. These are related to its semantics. In this context, the goals of sympérasmological studies are: (i) capturing (decoding) meaning, (ii) interpretation of meaning by systems, (iii) aggregating different meanings, and (iv) consequences of meanings. There is a lack of underpinning theories. Currently, it is deemed a rational position that digital computers are not (and perhaps will never be) able to understand the meaning of any body of knowledge. It is also widely debated if intellectualized systems can or not behave intelligently without understanding meaning. Not excluding these, semantic studies may intend to explore how knowledge of systems conveys meanings in different situations and contexts, how processing of meanings is possible, and how to activate meanings in decision making. Management of meaning is also associated with representational issues. Though there are automatically generated semantic knowledge graphs, the currently used knowledge graphs need human interpretation.

Since application problems should be solved by the available knowledge, systems knowledge induces pragmatics. In the framework of sympérasmology, pragmatics is the study of practical aspects of system operation, problem solving, goal attainment, and in particular the role of synthetic knowledge in these. The major concern of sympérasmological pragmatics is what ways knowledge can be used efficiently or optimally in various contexts of problem solving. It deals with these issues based on theoretical rather than practical considerations. Pragmatics of system-level knowledge is a kind of meta-semantics, which informs about the problem solving potentials in application contexts. Various pragmatic relations are to be investigated, such as adequacy, modality, parsimony, handover, correspondence, and dialogue relations, which are usually contextual and agent-dependent.

Like sympérasmological pragmatics, sympérasmological apobetics is concerned with relations, which are about the effects of attainment and not about the ways of attaining a goal. Apobetics discloses and explains relations from a purposive, goal, effect, association, implication, and/or feeling aspects. Apobetic studies of sympérasmology are interested in the outcomes of realization of an intended operation based on given bodies of knowledge. It also studies the reflections triggered by successful or unsuccessful problem solving and the role played by the structure and quality of the synthetic knowledge in that. The studied relations are such as (i) corporeal, (ii) perceptive, (iii) cognitive, (iv) affective, and (v) combined apobetic relations.

The multi-faceted phenomenon of system-level intelligence is waiting for clear definitions and consolidation from sympérasmology. Whereas an exact definition of its possible gradations (degrees of intellectualization, such as smart, cognizant, intelligent) is a taxonomical opportunity, providing a theory of optimal problem-solving intelligence is a basic practical necessity. In the quest for clarifying system-level intelligence, this domain of sympérasmological studies largely overlaps and complements artificial intelligence research and development, but it also has its legacy due to its focus on the probabilities and possibilities of system-level knowledge with or without system consciousness, cognizance, feelings, emotions, instincts, and values. Furthermore, the relationship of system-level intelligence and (unsupervised) automation also lends itself to useful sympérasmological studies. In this context, the notion of ‘artificial wisdom’ deserves attention, which is deemed indispensable for turning externally provided control to internally provided supervision, as the basis of automation.

As the ability to act critically or practically in any given situation, artificial wisdom shifts the focus from ‘knowing what’ and ‘knowing how’ to’ knowing why’, and by doing so, it leads to some sort of meta-knowledge concerning the overall operation of intellectualized engineered systems. Meta-knowledge involves complex logical, pragmatic, ethical, etc. judgments, decision-making, and anticipating (seeing beyond the present). As the literature evidences, current research efforts aim at replacing system wisdom by meta-knowledge constructs, which may remain external (separated from the synthetic knowledge) and internal (combined with synthetic knowledge).

5.9. Faculties of synthetic system knowledge

As a study domain in sympérasmology, the plural noun, 'faculties', has been used to express the whole (total) of inherent, acquired or engineered intellectual or physical, generic or specific features, abilities, capacities, and power that synthetic knowledge and system intelligence, respectively, own. The permanency of faculties is not supposed. Attributes are intrinsic, constitutive, and substantial qualities, features, or abilities of an entity (but are not its functional/structural parts). Holsapple (2004) defined knowledge attributes as dimensions along which different instances of knowledge can vary. An attribute dimension comprises a range of categorical and nominal values. Categorical values are, for instance, names of types of knowledge, and these may form multi-level taxonomies. The total of the attribute dimensions forms an attribute space. Every point of this space represents one particular instance of knowledge and disposition, which leads to abilities under certain conditions. Key attributes are unique, exclusive, and differentiating characteristics of an entity (such as a unique name or identifier). In practice, a combination of them can uniquely describe (and thus differentiate) an entity. Usually, a body of system knowledge can be characterized by a large number of shared attributes, while the key attributes are limited in terms of their cardinality. They form a differentiating attribute profile.

Manifestation of knowledge is about how the problem solving power of knowledge is conceptually rendered, and not about its alternative formal or informal representations. In an abstract way, manifestation can be understood as 'embodiment for purpose', which rests on using different realization principles. Manifestations are realized by some process of logical, conceptual, semantical, etc. modelling Aamodt, A. (1995). The sympérasmological study of knowledge manifestations focuses on how the (problem solving) power of knowledge is captured and operationalized in a specific manifestation (a concrete model). It may identify general principles for establishing the corpus of knowledge, as an extensive large collection of semantically interrelated knowledge elements. Thus, from a practical point of view, knowledge manifestations can be viewed as conceptual constructs (and formalized as semantic frameworks) of computational representations of that part of the real world which the concerned system is to reason about. A manifestation of knowledge may render a range of pertinent knowledge models that are fit for purpose.

From a computational viewpoint, system-level intelligence is the result of synergistic interoperation of bodies of synthetic knowledge and the associated processing mechanisms. The relationships among synthetic knowledge, processing mechanisms, and application problems lend themselves to sympérasmological studies naturally. A (problem-solving) mechanism is behind a purposefully arranged set of computational algorithms and a chosen set of chunks of knowledge. In this disposition, the notion of mechanisms follows the classical interpretation of Descartes as operational essences of the physical world. Implicit mechanisms are sought to explain how a phenomenon is produced, while explicit mechanisms are created to facilitate carrying out some tasks. By studying the relationship of synthetic knowledge and the intellectualized operations of systems, sympérasmology intends to understand the causal and operational synergy of the system constituents. As a sort of cognitive equipment for knowledge processing, both non-ampliative and ampliative mechanisms have been studied recently.

Handling of knowledge is traditionally a task of knowledge engineering and knowledge management. However, they handle knowledge at statistical, syntactic, and partially semantic levels. Knowledge engineering considers the objective and structure of the problems to identify how a solution or decision can be reached through knowledge processing, whereas knowledge management focuses on the process of defining, structuring, retaining, and sharing knowledge and experience of humans in order to improve efficiency and to save resources. The interest of sympérasmology is not in pre-processing, representation, and distribution of human knowledge, but in treating it as an intellectual artefactual object. The theoretical and conceptual perspectives on handling synthetic knowledge establish two interplaying viewing windows: (i) the relevance of the various handling operations to a particular synthetic knowledge, and (ii) the influence of the nature of the synthetic knowledge on the handling operations. As an example, sympérasmological studies are supposed to find answers to questions such as: 'What criteria should raw system knowledge fulfil?' and 'What correspondence, coherence, and consistency conditions are to be satisfied?' These show that the concept of handling is more a cognitive act, than a technological one for sympérasmology.

Likewise, system supervision may appear as a sympérasmological topic, and not only as a computational, organizational, or social one. What underpins the importance of sympérasmological studies in this direction is the growing intellectualization level of engineered systems that changed their relationship to humans as well as to other systems. This domain of interest has the human–system (or system-human) relationship in the focus and investigates it from the perspective of the increasing synthetic knowledge of intellectualized systems. New issues concerning these systems originate in that they: (i) operate as alert and proactive actors in multiple contexts, (ii) initiate contacts in various forms with human stakeholders, (iii) interact with the embedding environments and the tasks, (iv) obtain control information from real life processes, (v) perform certain level of functional, structural, and/or behavioral self-adaptation, (vi) alter the goal of operation towards a more favoring ones, (vii) establishing social relations within and outside the embedding environment, and (viii) show personalized behavior towards other natural and/or artificial actors. Handling these needs dedicated self-supervision knowledge, which differs, but is not completely separable from problem-solving knowledge.

5.10. Implications of synthetic system knowledge

The general term 'implications' is used to refer to capturing consequences which are not explicitly stated or suggested, but can be drawn as cause-effect relations from knowing the phenomenon (synthetic system knowledge). Implications express the possible significance of the phenomenon, the need to do something, or the effect that an action or decision will have on something else in the future. Thus, the consequence can be a need for action or for a state to get involved in.

The word 'potentials' concerns the total of the inherent physical, cognitive, perceptive, and emotional capabilities that can be operationalized by a system to be able to execute problem-solving actions. In terms of problem-solving, these together are often referred to as 'power of a system', and understood as a permanently available but varying characteristic. As an example: a neural network always learns a model (even if it is incorrect with regard to the features of data), but the system has the potential to improve the learnt model when it is a subject of refined training. System potentials are relative to some target problems and may include (i) abilities, (ii) affordances, and (iii) dispositions. For each of them, a stimulus condition (as a causal relationship) is needed to exhibit its power, as potentials, abilities that may exist even when they are not concretely displayed. Affordances are those possibilities that are recognized by a system in the process of problem solving (in spite of the fact that they are not predefined). Disposition carries the tendency of a system to act in a certain manner under given circumstances (e.g. the disposition of ice to melt when heated by some means). In principle, the 'will of action' can be viewed as a proper form of potentials, as long as it is interpreted as an independent decision of an agent in the course of performing actions.

Dependability is the overall quality of realizing the defined functions (operations and services) of a system. From the perspective of sympérasmology, dependability is projected onto the synthetic knowledge of a system. It predicts and measures how reliably and persistently an assigned body of knowledge can provide proper solutions for target problems. As a complex indicator, dependability also concerns the reasoning mechanisms working on the synthetic knowledge. Dependability may be expressed in terms of a limited set of specific, context-dependent indicators such as (i) availability, (ii) resilience, (iii) controllability, (iv) maintainability, (v) performance, and (vi) accuracy. Dependability is of crucial importance in the case of self-adaptive intellectualized systems, where problem-solving knowledge intertwines with self-adaptation knowledge.

Complex intellectualized systems become compositional through their problem-solving knowledge. At system level, this should be a holistic whole, rather than an aggregate of discrete chunks of knowledge. Compositionality of knowledge structures and inferring procedures is largely analogous to linguistic compositionality. Tani (2014) has shown that compositionality is necessary for higher order cognitive tasks and can be acquired by means of self-organizing hierarchical dynamic structures. In essence, compositionality is about combinatorial manipulations of concepts, objects, actions, plans, and problems in order to synthesize consistent and dependable knowledge through cognitive mechanisms.

As a fundamental concept, 'norms' are measures first created in the social sciences. In general, norms are accepted or acceptable standard ways or qualities of behaving or doing things. Common sense incorporates implicit norms which go with the very use of such notions as 'belief', 'knowledge' or 'judgment'. Sympérasmology should create and assess norms in an explicit and reflexive way, and should evaluate system intelligence and SSK in light of these norms. From the perspective of sympérasmological studies, a norm may express both (i) the kind of generic expectation to be reached in each individual case, and (ii) a designated or wishful level to act accordingly on knowledge level. Sympérasmological norms may be prescriptive (encouraging positive system operation), or proscriptive (discouraging negative system operation). There are many categories of norms that are relevant in the context of IESs and need to be considered in sympérasmological studies. Some of these (in alphabetical order) are: (i) cognitive norms (quality of problem solving potential), (ii) ecological norms (measure and rules of environmental impacts), (iii) ethical norms (the impacts of operation and fulfilment of moral expectations), (iv) practical norms (latent informal, de facto, or tacit norms underlying the practices), (v) social norms (societal rules or expectations for contextual behavior), and (vi) technological norms (principles of right problem solving operations

Advancement is one of the most intricate and challenging concepts for comparing two states of an intellectualized system. Advancement of synthetic knowledge concerns qualitative rather than only quantitative characteristics. It is more obvious to trace past advancements based on retrospective analysis, than expectable advancement that needs analysis of the overall scientific, technological, social, etc. trends as well as generating realistic visions. The central questions for sympérasmological investigations are whether (i) the conventional separation of 'human-imitating systems' and 'an-ideal-rationality' oriented systems will exist, (ii) the demarcation line between 'reasoning-centered systems' ("think like humans") and 'behavior-centered systems' ("act like humans") continues to fade away, (iii) will a symbiosis be formed between human intelligence and artificial intelligence, or (iv) the 'intelligence explosion' reaches a state ('superintelligence') that goes beyond human supervision in the future. These may imply radical changes in the kind of system knowledge, with the possibility of not generating and using anthropometric models.

5.11. Reflections and open issues

Sympérasmology has been proposed as an investigation domain and theory of epsilon-knowledge, which is represented by the ever-growing, complexifying, and becoming-independent SSK. Its fundamental concepts and principles have their roots in the very nature and procession mechanisms of SSK, and system-level intelligence in general. It is supposed to complement gnoseology and epistemology. Both the proposed conceptual framework, and the range, scope, and the domains and approaches of the investigation need further elaboration. Sympérasmology has the potential to make genuine and useful theoretical, methodological, and praxiological contributions different to those of systems science, technology management, and systems engineering. The release of the idea of sympérasmology for a broad public debate and encouraging all interested scientists, engineers, philosophers, and practitioners to address the proposed or any other relevant issues and questions may create its own literature. The benefits of dealing with sympérasmology are not only in a clearer academic view and a comprehensive underpinning theory, but also in the opportunity of more dependable innovation strategies and better engineering decisions about proper intellectualization of systems for industrial and social applications. In follow-up research and debates, the following issues need to be addressed:

- Formulation of a philosophically correct designation and specification of the objectives, interest, cognitive engine, and approaches.
- Dispositioning with regards to ontology and methodology as branches of science philosophy, as well as to gnoseology and epistemology.
- Consolidation of the spectrum of the relevant domains of interest and further articulation of an indicative study program.
- Acknowledging the principle of inseparability and strategizing the refinement of dealing with the whole (and not only the two interrelated constituents) of system-level intelligence.
- Deepening the knowledge concerning all intrinsic domains of interest with a view to their dependency, causality, as well as to their ampliative and implicative relationships.
- Establishing quality and dependability criteria for sympérasmological investigations and their outcomes.

Chapter 6

6. Enabling prognostic systems thinking

6.1. Research objectives and approach

System thinking is a holistic approach to analyzing how the constituent parts of a complicated system interrelate and how the system works over time and within the context of larger systems (Checkland, 2000). The current propagation of intellectualized and socially embedded engineered systems raises many novel and crucial technological, managerial, social, organizational, business, environmental, human, and so forth, issues. It has been recognized that traditional systems thinking is not able to consider many new aspects and to provide reliable forecasts on accelerated developments. It has been argued in the literature that there is a need for **prognostic systems thinking** (PST) that considers not only the observable manifestation of systems, but also the trends of their overall development. It is also argued in the literature that there is a growing need for prognostic normative reasoning (Oh et al., 2013). Based on a comprehensive literature study and concept relationship analysis, the research reported in this chapter: (i) critically analyzed the status of systems thinking, (ii) identified new influencing factors, and (iii) rendered a semantic network as a basis of prognostic reasoning. In line with other published efforts, the ultimate goal of the study was to derive a comprehensive set of investigational concerns on the basis of an extended conceptual framework.

This chapter (i) explains the reasons for the aforementioned need, (ii) identifies the pillars of contemporary **analytic systems thinking** (AST) as well as a set of up-and-coming new pillars, (iii) presents a first iteration of the novel conceptual framework, (iv) discusses the proposed investigation concerns for prognostic systemic analysis, and (iv) casts light on their implications with regard to prognostic system thinking. Furthermore, some currently open issues (such as determination of the boundaries of systems, completeness/sufficiency of a given set of concerns, and ranking/preference of concerns) and research opportunities (such as methodological support of deriving concerns, objective concerns assessment in various application contexts, and computer support of PST) are suggested.

The inspiration for this work came from the recognition that it had always been the best practice in natural sciences, especially in physics, to reconsider the existing theories time-to-time in light of the progression, and to take into account the emerging paradigms that generated new insights. The concrete goal was stated as to take a step towards the indispensable enablers (principles, frameworks, and guides) of PST that assume the identification and interrelation of the essential characteristics of next-generation intellectualized systems as well as the related societal issues. Some of these are addressed - among others - in (Straus, 2021) and (Waddock, 2016). As a starting point, it was assumed that the **prognostic conceptual framework** should be based on an adapted conceptual framework that considers (i) the shifting system paradigms, (ii) the outcomes, interplays, and implications of the trends of system realization, and (iii) the different humans-systems and systems-systems relationships (Hammond, 2005). Rendered as a semantic network, this conceptual framework rests on a set of foundational concepts (semantically interconnected pillars). The proposed conceptual framework rests on a limited set of foundational concepts (semantically interconnected pillars) and their joint implications. The reported research is based on a comprehensive literature study, concept relationship analysis, and mind mapping. The contents of this chapter have been compiled from the following peer-reviewed publications: H16 and H17 (see Appendix A.1.1).

6.2. Fundamentals of analytic systems thinking

System thinking was introduced as a new paradigm of thinking about complexities and heterogeneities, opposing the reductionist thinking that is typically pursued in monodisciplinary sciences. For this reason, it is in juxtaposition with and complements scientific thinking, which imposes reductionism rather than holism in its view (Albright and Runehov, 2013). Historically, the roots of modern system thinking can be traced back to the 1950s. As discussed by Forrester (1990), an important milestone was when the Systems Dynamic Group at the Sloan School of Management of MIT was founded in 1956. There are many originators of systems thinking who offer new insights and principles, and many refiners who contribute to improving existing insights, definitions, and models (Monat and Gannon, 2015). Richmond (1991) defined the term ‘systems thinking’ as the art and science of making reliable inferences about behavior by developing an increasingly deep understanding of the underlying structure. Systems thinking has become the major enabler of complex (transdisciplinary) system analysis and assessment (Richmond, 1993).

A specific form of systems thinking is the investigation approach of traditional **reductionist system analysis**, which studies systems by breaking them down into separate elements (Zexian and Xuhui, 2010). System thinking is variously interpreted in the literature. For instance, as (i) an ontological platform, (ii) a methodological framework, (iii) a means of enhancement, (iv) a model of holistic comprehension, (v) a knowledge sharing mechanism, (vi) a set of disciplinary principles, and (vii) a construct of learnt competences. It has placed the emphasis on comprehensive, diagnostic, and critical reasoning about complex situations, which resulted in its overall analytic nature. Often, elements of design, societal, and computational thinking as well as pragmatism can be identified in systems thinking, which is a kind of indication of the overlaps of the various thinking paradigms.

Systems scientists have made diverse and divergent attempts to place the theory and methodology of systems thinking onto various philosophical platforms. Some of them have tried to provide a universal description of its goals and principles (Buckle Henning and Chen, 2012). Five major philosophical platforms of systems thinking have been identified in the related literature recently, namely (i) functionalist, (ii) interpretivist, (iii) postmodernist, (iv) emancipatorist, and (v) transdiciplinarist, as shown and exemplified in Figure 6.1. From the perspective of the theory and practice of operational research, (Jackson, 2009) provided a comprehensive historical overview of the functionalist, structuralist, and interpretivist applied systems thinking. Senge (2006) defined systems thinking as a specific discipline that offers a way of thinking about and understanding the effects and interrelationships that shape systems, and a language for describing their

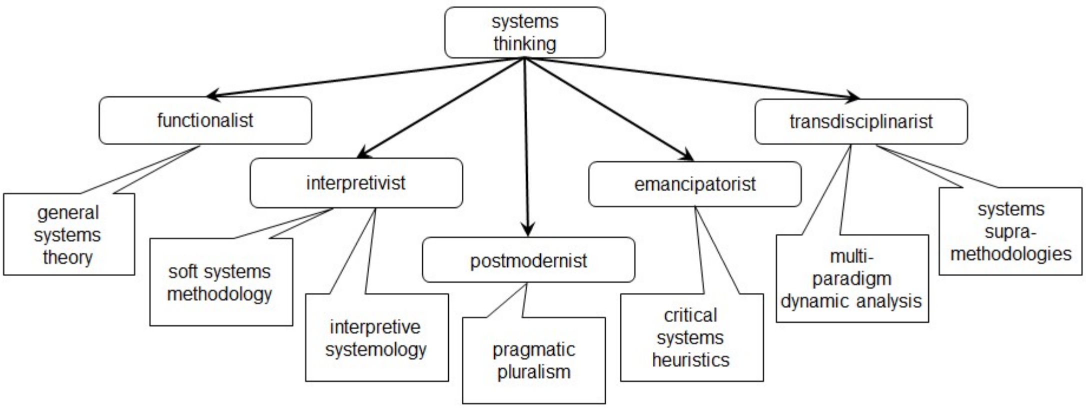

Figure 6.1: Philosophical platforms and main approaches of analytic systems thinking

behavior. Consequently, the precise meaning of systems thinking as a model of holistic and critical human thinking has remained ambiguous and has been characterized by Cabrera et al. (2015) as the field of a baffling array of methods and approaches.

As forecasted by the literature and also this dissertation, even near-future systems will essentially differ from those that were regarded as references at the time of making the first road-paving efforts to develop all-embracing system theories, frameworks, and methodologies (Fazey et al., 2020). The current popularity and proliferation of highly intellectualized and **socially deeply embedded engineered systems** raise many novel technological, managerial, social, organizational, business, environmental, human, and so forth, questions about the present and future of systems. It is strongly believed that the power of the systems paradigm can improve the way people exist and operate in the existing world (Steels and López de Mantaras, 2018). On the other hand, application of the traditional system thinking frameworks and principles for studying phenomena and situations related to emerging systems typically results in limitations and incompleteness (Ulrich, 2013).

In overall, the ontology and epistemology of contemporary systems thinking is a loose, broad, and eclectic collection of theories and methodologies, mirroring the divergence of systems engineering practice (Wan, 2011). The main assumptions of traditional systems engineering are: (i) absoluteness, (ii) unambiguity, (iii) sequentiality, (iv) rational actors, (v) reductionist, (vi) central controlling, (vii) static solution, (viii) mechanistic factors, (ix) deterministic behavior, (x) context independence (Pennock and Wade, 2015). Systems thinking has been regarded as a management discipline that concerns understanding of defined systems by examining the linkages and interactions between the components that comprise their entirety (Hürlimann, 2009). In the field of management science and practice, the view of **holistic systems thinkers** is sharply contrasted with the view of **event-oriented thinkers**. The latter thinkers assume that each event (or in other words, solving a related problem) has a specific cause and handling an event means finding the cause and fixing the problem according to that.

Keating and Gheorghe (2016) argued that analytic systems thinkers see a problem (the structure and internal/external interactions of the system) entirely differently, namely as the cause of any regular behavior or misbehavior, and use potential feedback loops to achieve the goal. Stave and Hopper (2007) argued that seven systems thinking characteristics are required and are sufficient to describe systems, which include (i) recognizing interconnections, (ii) understanding dynamic behavior, (iii) identifying feedback, (iv) differentiating types of variables and flows, (v) using conceptual models, (vi) creating simulation models, and (vii) devising testing policies. The contemporary professional literature does not offer a concrete methodology to apply these general principles to (evolving) intellectualized engineered systems.

As summarized above, a determining trend is that the number and kinds of engineered systems are rapidly growing, in particular, in the realm of intellectualized engineered systems (Wendt et al., 2009). New opportunities are offered by this, but also new challenges (uncertainties and incompleteness) are created for system science and engineering (Drack and Apfalter, 2007). Many of these challenges have been recognized and various efforts have been made to adapt the theory and practice of systems thinking to them (Mahmoudi et al., 2019). To understand the behavioral patterns that arise in systems of different systemic structures, the concepts and models of **system archetypes** were introduced (Kim & Anderson, 1998). The importance of system archetypes is in that they can be used both diagnostically (to gain insights into the behavioral structure of a system) and prospectively (to help foresee intended or unintended behavioral consequences) (Braun, 2002). The system archetypes concept needs to be updated with regard to intellectualized engineered systems.

In **investigative systems thinking**, a phenomenon to be explained is viewed as part of a larger whole, a 'system', and is explained in terms of its role in that 'system' (Daellenbach et al. 2012). The 'system' is the kernel of the process of understanding the real world (Jaradat and Keating, 2016) and has been applied, among many others, to the study of medical, environmental, political, economic, human resources, and educational systems. Behl and Ferreira (2014) clarified the difference between 'individual systems thinking' (that is based on the ability of an individual engineer to demonstrate systems thinking) and 'collaborative (team) systems thinking' (that is based on an emergent behavior resulting from the interactions of team members and utilizing a variety of thinking styles, design processes, tools and communication media to consider the system, its components and dynamics towards executing systems design.

6.3. Legacy of prognostic systems thinking

The above overview has cast light on four important issues (namely, objectives, characteristics, mechanisms, and competences), which need to be operationalized at investigating systems from a managerial, organizational, social, developmental, or another viewpoint. Notwithstanding the importance of these, we must ignore neither the dramatic changes, nor the implied issues that have been taking place (both in the conceptual realm of systems and in our daily life thinking) due to the human endeavor to create systems that reproduce parts of human intelligence in various forms, or to create intellectualized autonomous systems. Possessing system-level synthetic problem solving intellect, the latter systems rely on processing application-specific SSK by ampliative reasoning mechanisms. Currently, fast advancements can be witnessed in both mentioned domains of interest, culminating in knowledge and resource exchange among cooperating systems. It can be foreseen that the next generation of intellectualized systems will largely differ from the current socio-technical systems and move towards socialized and personalized smart CPSs. The difference will be not only in their (i) functional complexity, (ii) architectural heterogeneity, (iii) level of intellectualization, (iv) operational smartness/intellect, (v) natural/social/societal embedment, (vi) personal interrelationships, but also in the (vii) services, values, and experiences provided by them.

Dealing with systems equipped with system-level problem-solving intellect raises the need for different reasoning models and aspects of investigation. These kinds of systems are deemed to (i) operate collaboratively in an environment with other systems, (ii) possess cognitive abilities such as perception, action control, deliberative reasoning, or language use, (iii) follow behavioral principles based on rationality and social norms, and (iv) have the capacity to adapt through learning. These are why it is necessary to lay down the foundation of a **novel systems thinking** that will support not only reflective, but also prognostic problem assessment and solving. Future systems thinking is supposed to be proactive and to provide new strategies and blueprints for moving to a highly-intellectualized (perhaps intelligent) systems future. Due to the inability of the conventional approaches to address forthcoming challenges, prognostication is expected to become a new affordance of future systems thinking. That is, future ST models, frameworks, narratives, renderings, etc. are supposed to capture not only the (operational and managerial) complexity and problems associated with next-generation systems, but also the shifts of paradigms, the observed trends of changes, the varied manifestations, realizations, and interactions of systems.

As introduced above, PST differs from both predictive systems thinking and critical systems thinking (Baker, 2007).). The former one focuses on the phenomena expectable in a particular states of systems, whereas the latter one is regarded as a multi-methodology that combines methods and practices from various systems thinking domains, such as system dynamics, soft systems methodology, sociotechnical systems, and others, in order to better understand and address a recognized problem. It also differs from soft systems thinking, which pursue tackling unstructured problems (soft problems) to achieve improvement to the system through a multistage process of information gathering, description, analysis and debate. For instance, in the field of human activity systems (Burge, 2015).

Thus, the goal of my work has been to devise **prognostic analytic enablers** and skills considering the perpetual and accelerated changes of the paradigms and trends. My intension was to move towards a system thinking that takes into account all rapid changes that influence how we see, treat, and benefit from complex systems, no matter if they are natural, fully or partially engineered, or hybrid systems. I also intended to obtain new insights into systems thinking - closer to the state and the truth of today - by means of new interpretations and elaboration of a novel conceptual framework. These are the main ingredients of the novelty of my approach. Be that as it may, it is fair to remark that, in spite of its novelty, my work is not unique in terms of rethinking systems thinking. Other authors have made similar efforts, in particular in the field of social science (Gradinarov, 2015), but they typically started out from different assumptions or had different purposes.

In my interpretation, first of all, an **evidence-based conceptual framework** (CFW) is needed that can serve as a guide to (holistic) systems thinking scenarios and models, and that can facilitate prognostic analyses (without pretending being a technological roadmap to the world of intelligent systems). One of the requirements is that the conceptual framework should include ontological and methodological concepts that synergistically complement those discussed by Pickel (2007). In addition, the CFW should provide adequate view and knowledge on next-generation systems independent of their genre, and should make it possible to formulate relevant and accurate predictions about their manifestations, realizations, behaviors, and changes. Furthermore, the CFW should provide a high-level reasoning mechanism for PST. Obviously, the framework cannot be restricted to the paradigmatic features of one particular system. Instead, it may cover all systems that do not contradict the conceptual pillars that the CFW is based on. However, the intension resonates with that presented in Sun et al. (2014).

The main expectations for PST are as follows: A primary task is to help identify emergent patterns conducive to explanation and prediction of future systems, and not only to their systems theoretical/engineering description and identification. PST enables, even forces, the different observers to look beyond what is in front of them and to see what is probable and what is possible. PST complements the currently identified three levels of descriptive systems thinking, that is: (i) the basic level (involving the recognition of interconnections, identifying feedback, and understanding dynamic behavior), (ii) the intermediate level (adding the differentiation of flows and variables, and using conceptual models to the basic level), and (iii) the advanced level (including the creation of simulation models, and testing policies). It introduces three additional levels associated with a fundamental understanding of the patterns that characterize next-generation systems and with the explanation and prediction of their behavior under various circumstances. Furthermore, it also introduces novel system archetypes which help model the essence, operations, and implications of intellectualization of engineered systems. Actually, these archetypes are already captured implicitly in the conceptual framework that describes the paradigmatic characteristics as well as activities/behaviors.

6.4. Pillars of a conceptual framework

It is well-known from the literature that systems thinking revolves around a handful of concepts that anyone who is determined to learn can master, with study and practice. In the process of devising the proposed conceptual framework, the key concepts of analytic systems thinking have been reused as traditional conceptual pillars. The conceptual framework has been specified as a combination of the widely accepted conceptual pillars of traditional systems thinking and a finite set of the up-and-coming (emergent) conceptual pillars. The pillars have been formulated in an abstract and general manner in order to keep them applicable to different types of systems and to others. In the above conceptualization, the key elements of prognostic reasoning are not causal relations, but semantic relations of the conceptual pillars. The **traditional conceptual pillars** have been reused in the following interpretations:

boundedness	system has a domain of realization and operation defined by its individual boundaries
causality	system behavior is due to the transitive relationships of parts as influenced by the environment
compositionality	system as an entirety is more than the sum of its parts
consolidation	system changes towards a sustainable existence over time
distinctiveness	system has a purpose that defines its distinct functionality, manifestation, and/or realization
equilibrium	system's overall behavior is the result of triggering and balancing processes that may lead to stable or emergent phenomena
holism	systems should be thought about seeing the big picture rather than the parts of it
impact	systems with multiple outputs exert cumulative patterned effects on their embedding environment and stakeholders
interconnectedness	system is an arrangement of a finite number of purposefully interrelated parts

The above specifications reconfirm the fact that the traditional conceptual pillars of system thinking are about how systems manifest, rather than about how they evolve or are (self-)transformed over subsequent generations. They also disclose that traditional systems thinking neither focuses on evidential reasoning about possible and probable long-term developments, nor serves as an adaptive basis for reasoning about changing humans-systems relationships. These are supposed to be provided by additional pillars.

Altogether, eleven evolving concepts have been considered to capture the influence of specific trends, such as growing intellectualization, self-management of resources, increased autonomy, deep social embedding, and so forth. These non-forever-valid concepts establish the additional pillars of the extended conceptual framework. The sources of these conceptual pillars are varied, as well as the concerns implied by them (discussed in the next section). The sources of them are, for instance: (i) theories produced by empirical research, (ii) trends and implications of observed technological developments and societal demands, (iii) principles derived from and by philosophical speculations, (iv) postulates, conjectures, and assumptions concerning future situations, (v) economic conditions and projections, (vi) political initiatives and policies and (vii) subjective personal beliefs, inceptions, and opinions. In the specification of the **additional conceptual pillars**, these sources have been considered with a preference/emphasis indicated by the order of mention above. Thus, in an alphabetical order, the set of additional conceptual pillars of the proposed system thinking framework includes the following concepts:

annihilation	systems reflect the disappearance of thingness in terms of physicality of their constitutional entities

complementariness	systems complement delimited humans and humans complement delimited systems
exploitation	reasoning mechanisms and synthetic knowledge of systems aggregate into a shared industrial asset
ineptitude	systems are constrained in terms of replication of genuine human abilities such as cognition, abstraction, feeling, and affection
inseparableness	systems may not entangle their purpose of existence, behavior, and performance from human will and meaning
intelligence	intellectualization of systems extends the cognitive capabilities concurrently in the perceptive, reasoning, and communicative domains
naturalization	intellectualization, organization, socialization, and personalization of systems are coexisting dimensions of progression
organization	systems are interconnected with each other in arrangements like systems, swarms, and/or societies of systems
personalization	functional capabilities of systems are extended with human behavioral traits and para-functional abilities
self-management	systems are capable of purposefully changing themselves towards improved overall performance based on reflective learning
socialization	social embedment of systems may range from socially interacting to socially behaving systems

The total of the traditional pillars and the additional pillars are graphically shown in Figure 6.2. The former ones are drawn by thick lines, and the latter ones by thinner lines. They together identify the primary attention points for prognostic systems studies. Moreover, they help analyze the changes in terms of their manifestations and impacts on systems in a holistic way – even over their subsequent generations (that is, beyond the life cycle of a particular system).

A pertinent question has been how the compiled set of conceptual pillars can be transferred into concrete interrogative statements for the targeted systematized inquiries? It has been hypothesized that the conceptual pillars individually, or their pairwise or multiple relations, imply one or more chunks of information for the investigations. These chunks of information have been called **concerns** and textually formulated as arguable statements. Eventually, the conceptual framework has been constructed based on the semantic relationships of the particular pillars and interpreting the implications of their interplay. It facilitates follow-up knowledge exploration, chunking, and association, and eventually leads to a semantic network.

6.5. Framework as a network of semantic relationships

In the literature, framing the knowledge constructs associated with systems thinking and the development of logical frameworks are widely addressed issues. By definition, a framework is a construct of assumptions, concepts, values, and/or practices that represents a part of the existing or imaginary reality and constitutes a specific way of dealing with it. A framework holds chunks of knowledge logically and semantically together for a particular purpose. The most frequently occurring types of frameworks are (i) conceptual (logical), (ii) taxonomic (architectural), (iii) methodological (procedural), and (iv) practice oriented (model-based) frameworks (Tepjit et al., 2019). An agreement seems to exist that a conceptual framework is a logical arrangement of a set of physically or theoretically supported concepts and values that constitutes a way of viewing reality and serves as a frame of reference. It lends itself to making conceptual distinctions and to organizing a set

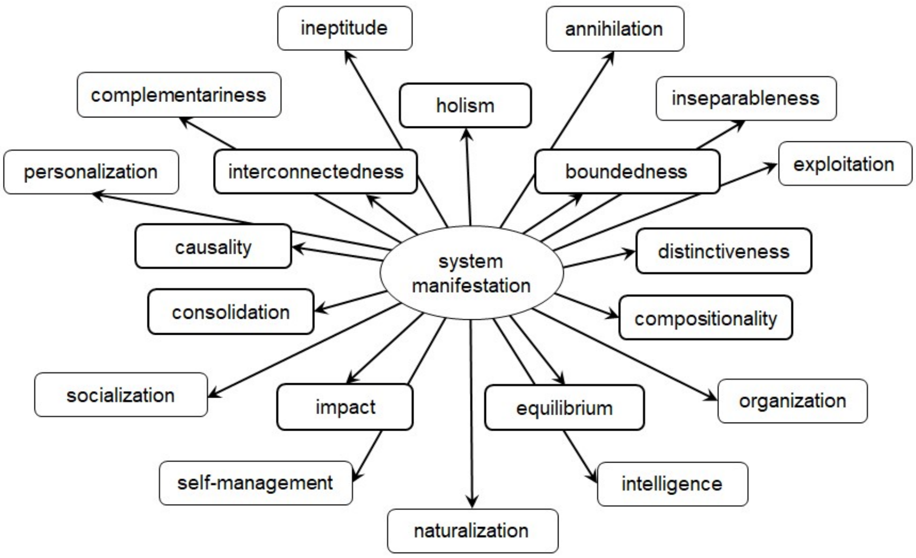

Figure 6.2: The traditional and the additional pillars combined with the traditional pillars of systems thinking

of concepts according to their semantic relationships.

The pillars can be sorted into three sub-groups, which carry information about three influential factors: (i) recent trends of progression, (ii) paradigmatic features of (intellectualized) systems, and (iii) relationships with humans. The conceptual framework for PST has been specified as a logical construct defined by the semantic relationships between pairs of pillars. A visual rendering of the **semantic network** of the proposed conceptual framework is shown in Figure 6.3. Implicative semantic relations may connect (i) two traditional pillars, (ii) a traditional pillar and an emergent pillar, and (iii) two emergent pillars. They do not have any importance over each other. Each relation may imply an arbitrary number of concerns according to the complexity of the interrelated concepts. The derived concerns are actually the means of operationalization of the proposed conceptual framework in PST.

The whole of the conceptual framework reflects an abstract view on the factors that are to be considered in current and near-future systems thinking. In the analysis, the concerns are projected onto the concrete situations that are considered as complicated systems. In practice, the projection means formulation of inquiry questions. This method supports

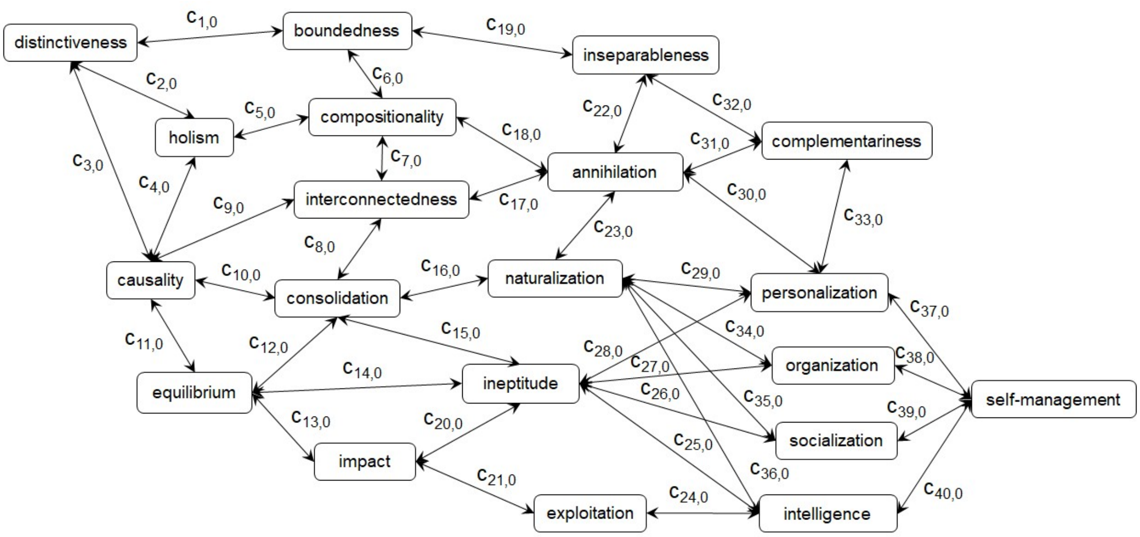

Figure 6.3: The proposed conceptual framework as a semantic network

forward-looking projections and reasoning prognostically over multiple evolving stages of systems. The trustworthiness of the prognostication is linearly proportional to the amount of the used data, information, rules, and human experiences.

6.6. Concerns-driven prognostic assessment

The term '**concern**' has been used to refer to various essential thoughts, matters of interest, and important issues related to the conceptual pillars and to their interplay. Concerns have been explored and formulated based on extensive literature study and semantic concept analysis. There have been different numbers of concerns identified about the various pillars. A challenge was to define the concerns on the same abstraction and comprehensiveness level. In the lack of objective measures or comparison methods, this was based on subjective judgment.

The concerns serve as guides to developing sets of questions to collect and analyze information for PST and deriving up-to-date mental and action models. In a somewhat different context, a similar work was done by Dorani et al. (2015). Depending on the system to be dealt with, it may be sufficient to consider a subset of the concerns, but there might be a need for the consideration of more or more articulated concerns in the case of a low-level comprehensive analysis. Obviously, prognostication becomes more factual when a large set of all relevant concerns are taken into account. For each edge of the semantic network, one or more concrete questions can be stated, based on which a concrete systems thinking model can be developed. On the other hand, system assessment and decision making eventually remain dependent on human views and interpretations.

To impose a simple structuring on the pile of the identified concerns, I have sorted them into three categories according to the classification of the pillars explained above. The categories of concerns are associated with (i) the traditional conceptual pillars and their interactions, (ii) the interrelated traditional and emergent conceptual pillars and their interactions, and (iii) the additional conceptual pillars and their interactions. In the background research, the exploration of all important concerns came along as an unexpectedly tough nut to crack due to the wide range of engineered systems and possible interests of human stakeholders. It is clear that explaining the concerns with one or more concrete practical examples would be helpful for the reader of the dissertation. Notwithstanding, this could not be achieved due to the obvious page limitation. Nevertheless, it is believed that connecting the presented textual formulation of the concerns to real life examples will not be challenging for systems researchers. Even a complete overview of the first inventory of concerns, which has been published in the article underpinning this chapter, could not be included here. Below, only representative examples of this extensive set of concerns are given.

Representative concerns associated with the traditional conceptual pillars and their interactions are:

$C_{03,0}$: distinctiveness <==> causality:

- $C_{03,1}$: local and remote interconnections and interplays of the constituents and parts as a reason for the observed operation and manifestation of the behavior of a system
- $C_{03,2}$: direct and indirect influence of the environment as a reason for the observed operation and manifestation of the behavior of a system

$C_{07,0}$: compositionality <==> interconnectedness:

- $C_{07,1}$: methods and overheads of how new constituents can connect to and connected constituents can leave a system of higher-level arrangement
- $C_{07,2}$: methods and overheads of how an open boundary system of higher-level

arrangement can handle operational and structural dynamics

$C_{11,0}$: causality <==> equilibrium:

$C_{11,2}$: lasting deterministic behavior (i.e., the next state of the system and the environment is not or not completely determined by their current states and the completed actions)

$C_{11,2}$: lasting stochastic behavior (i.e., the next state of the system and the environment is not or not completely determined by their current states and the completed actions)

Representative concerns associated with the interrelated traditional and emergent conceptual pillars and their interactions are:

$C_{14,0}$: equilibrium <==> ineptitude:

$C_{14,1}$: incapability for adaptation due to the quasi-equal effects of reinforcing and balancing factors and processes

$C_{14,2}$: possibility of pre-emptive reaction to state changes before they occur based on forecasting the changes in system level

$C_{16,0}$: consolidation <==> naturalization:

$C_{16,1}$: compliance with the AI control problem (i.e. acknowledgment of the problematics of building AI-enabled systems that aid rather than harm their creators)

$C_{16,2}$: corrective self-reflection based on information about the effectiveness of the used computational approaches, the system's own overall behavior, and the historical successes and failures in solving application problems

$C_{20,0}$: ineptitude <==> impact:

$C_{20,1}$: exertion of cumulative, patterned effects by a system having multiple varying outputs for its embedding environment and stakeholders

$C_{20,2}$: indicators of the limitations and constrains of a system with regard to replication/reproduction of genuine human traits such as thoughts, abstraction, feelings, and affection

$C_{20,3}$: capacity of an intellectualized system and its constituents to have computational (artificial) sentience for positive and negative experiences

Lastly, representative concerns associated with the additional conceptual pillars and their interactions are:

$C_{24,0}$: exploitation <==> intelligence:

$C_{24,1}$: accelerating and decelerating factors (barriers and drivers) of adoption of a type of intellectualized system in various sectors

$C_{24,2}$: importance and convincingness of intellectualized (or intelligent) system characteristics from an investment point of view

$C_{24,3}$: relationship between the achieved level of intellectualization and the socially demanded level of intellectualization

$C_{25,0}$: ineptitude <==> intelligence:

$C_{25,1}$: limits or deficiencies of implementation and aggregation of overall intelligence in an engineered system and synthesizing system intelligence on higher-level system arrangements

$C_{25,2}$: limits or deficiencies of replication of cognitive abilities such as cognition, abstraction, feeling, and affection

$C_{25,3}$: potentials and impacts of a system that is capable or has no capability of supervised, semi-supervised, and/or unsupervised learning

$C_{32,0}$: inseparableness <==> complementariness:

$C_{32,1}$: measure of the system's insistence on or deviating from the human-stated overall purpose in the case of or towards (more favoring) opportunities and affordances

$C_{32,2}$: cooperation of systems that are not organized for success in the same way or are (or not) equally committed to successful goal achievement

$C_{32,3}$: guaranties that the system behaves as a robust (stable) transition system even under heavily dynamic operational and environmental circumstances

$C_{39,0}$: self-management <==> socialization:

$C_{39,1}$: ability to bridge the cognitive-social-technical gap between what and how the self-managing intellectualized system does and what a segment of society wants

$C_{39,2}$: involvement of social components such as human, culture, organization, context of use, usefulness, policies, and regulations by a self-managing intellectualized system

$C_{39,3}$: principles, norms, rules, laws, and ethics of socialization of self-adaptive, self-evolving, and self-replicative intellectualized systems

6.7. Reflections and open issues

Systems thinking has been proposed to extend our mental models to render comprehensive, whole-system perspectives. Thinking about systems is like viewing a (double-sided) coin. On the obverse side of the coin is the question about what constitutes systems thinking, whereas on the reverse side is the question about what enablers, knowledge, and capabilities a systems thinker should have in order to be efficient. The contribution of this chapter can be placed on the edge of the coin because it connects the principles of next-generation systems to the knowledge that system analysts should be aware of. The pillars and concerns have been interpreted from a pragmatist standpoint. This has its legitimacy because, as Gradinarov (2015) argued: “To a great extent, systematic thinking is constructivist in nature, as is modelled not by specific preliminarily set and axiomatically adopted basic principles but is dependent on the subject and must include a preliminary analytical stage”. This degree of freedom plays an important role in converting the reasoning model into PST models serving different applications and contexts.

Prognostic systems thinking is seen in this work as an investigative thought process that is framed according to the particular subject matter, purpose, and context. Approaches of traditional systems thinking assume permanency and slow overall changes of the investigated systems, and look at them in a retrospective way or in a snapshot. When characteristics and manifestations of systems evolve rapidly, the need for PST emerges. The objective of the presented research was to (i) identify concepts that characterize both static and dynamic systems, (ii) construct a framework that semantically associate these concepts, and (iii) explore topics (issues) for a prognostic conceptualization or investigation of evolving systems. The trends and implications of their developments have been taken into account for a broad spectrum of intellectualized systems.

- PST enables a systematic examination and validation throughout all stages of activities from the formulation of problems, through the implementation of solutions, to the evaluation of outcomes. The proposed method of analysis assumes evidence-based systematic reasoning and an objective decision-making approach.
- A prognostic investigation of a system includes (i) deriving arguable goals and assumptions, (ii) specification of the set of pertinent concerns, (iii) compiling expressive

sets of indicators and measures, (iv) determining the extent of the investigated system, (v) comprehensive examination of the concerns, (vi) discerning influencing factors and their effects/implications, (vii) linking and evaluation of the findings, and (viii) reflections and projecting out propositions.

- The completed research could provide only a first inventory of the possible concerns of analysis and obviously further studies are needed towards its completion, consolidation, and validation. Though the proposal allows seeing intellectualized systems and their evolution processes as a whole, it does not offer support for handling complexities.
- Development of computational tools may reduce the challenge and efforts needed. The conceptual frameworks and analysis models of traditional systems thinking are criticized for their (i) static, (ii) descriptive, (iii) not ampliative, and (iv) non-resilient character. The proposed extendable conceptual framework and set of concerns can be extended as needed by the near-future state of system intelligence, autonomy, technology, and socialization.
- Handling the hottest issues, such as (i) the growing functional automation and naturalization of systems in terms of organization, behavior, and socialization, (ii) self-management of personalization, organization, socialization, and cognition, and (iii) exploration of their effects on humans, organizations, and society as a whole, needs further research efforts.
- Discussed in recent literature, traditional systems thinking does not have a structured 'semantic language' that can harmonize thinking and facilitate communications across the various domains of natural, social, and (intellectualized) technical systems. The concept of pillars and derivation of concerns are a step forward in this direction. This also strengthens the transdisciplinary character of PST. Nevertheless, there is a likely need for extension and further refinement since there have been several open theoretical and methodological issues recognized.

Chapter 7

7. Aggregation and utilization of synthetic system intellect as an industrial asset

7.1. Research objectives and approach

Knowledge is one of the rare resources that becomes **more if shared**. That is why aggregation, combination, and distribution of individual (tacit) and general (scientific) human knowledge have received so much attention over the centuries. Schools, libraries, and conferences, as well as the Internet and digital warehouses make knowledge accessible for everyone. Can something similar happen with SSK? Can intellectualized systems increase the epsilon-knowledge by aggregating, blending, and reusing SSK? By what approaches and resources can this goal be achieved? Can system-level knowledge eventually manifest as a profitable industrial asset? These were the questions stimulating the background research, the latest results of which are presented in this chapter.

Chapters 3 and 4 presented the trends of intellectualization of CPSs and argued that system knowledge and reasoning mechanisms are the most essential enablers, which make such systems able to solve application problems and to maintain their efficient operation. This chapter looks into the near future. It claims that SSI can be converted to a new industrial asset and utilized as such. In comparison with the state of exploitation of the various genres of human knowledge, the exploitation of epsilon-knowledge as an industrial asset is still in an embryonic stage, but forward-thinking researchers and managers have already cast light on some opportunities, challenges, and benefits. Unfortunately, no overall theory of synthetic system intelligence exists and its conceptual framework, management strategy, and computational methodologies are still in a premature stage. These are the main reasons why no significant progress has been achieved in this field, contrary to the latent potentials.

This chapter is intended to contribute to the understanding of the roots of the system intelligence transfer problem and places system intelligence in the position of a new industrial asset. The presented material is based on a critical literature study and many formal and informal discussions with researchers. Various technological options for **transferring system intellect** to several other systems have been scrutinized. In fact, four families of analogical approaches to SSI transfer are briefly analyzed: (i) knowledge transfer based on repositories, (ii) transfer among agents, (iii) transfer of learning resources, and (iv) transfer by emerging approaches. They are seen as starting points of the development of dedicated computational technologies and management approaches. An attempt was made to include these technologies in a technological framework that is able to cover the heterogeneity of the systems and their system-level intelligence. The last section presents this overall procedural framework and discusses various issues of provisioning synthetic system intelligence as an industrial asset. The procedural framework identifies the generic functionalities needed for a quasi-autonomous handling of synthetic systems intellect as an industrial asset.

As discussed in Section 6.4, system intelligence is a complex problem-solving and state management power that is based on representation of factual data, information, and knowledge, and goal and context-dependent computational reasoning. It has a human-created initial part and a system-produced (becoming dominating over time) part, which has been referred to as **synthetic system intellect** (SSI). The initial intelligence of systems is

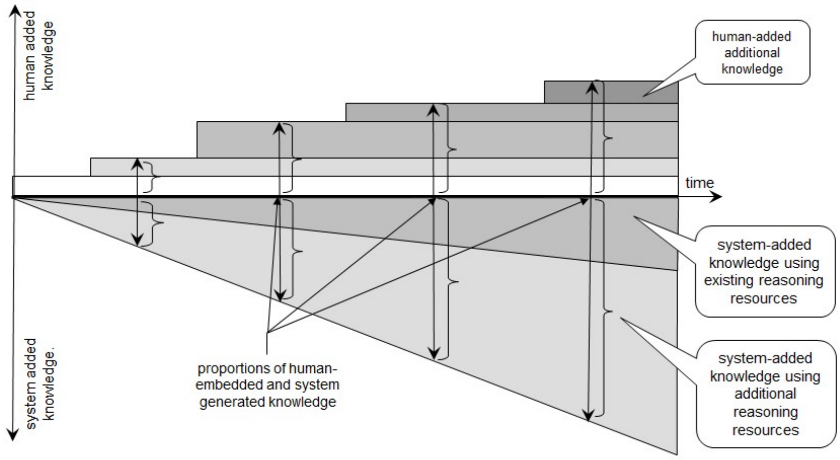

Figure 7.1: Overall changes in the proportion of the human embedded and the self-produced knowledge of intellectualized systems

usually incrementally modified or supplemented according to the foreseeable or actually experienced changes to operational conditions by human knowledge engineers.

However, autonomous systems are supposed to adapt their intelligence automatically under such circumstances or when, due to new goals, the elements of the initial intelligence may become outdated. What it means is that this part of the system intelligence can become not only significantly supplemented, but also partially or completely replaced by an evolving part that is self-acquired or self-generated by the system over its useful lifetime. The principle of this phenomenon is illustrated in Figure 7.1. The issue is how this autonomously managed synthetic part of system intelligence can be converted into a new industrial asset and utilized as such in an across-systems manner. Similar attempts were also made in the realm of exchanging human knowledge at enterprise level some 50 years ago (O'Leary, 1998), and in the realm of data and information management in the last three decades (Collins and Smith, 2006). This chapter emphasizes that time has come for the needed systematic studies and technology development in the context of aggregation, compilation, fusion, transfer, and reuse of the constituents of SSI. The content of this chapter has been compiled from the following peer reviewed publications: H12, H15, and H18 (see Appendix A.1.1).

7.2. Roots of the system intellect transfer problem

The origins of transferring various forms and resources of system intellect can be traced back to the 1970s. This was the time when it was recognized that system intelligence could be a **problem-solving power**. While the transfer of symbolic knowledge possessed by knowledge-based systems was in the center in the 1970s, the transfer of learning resources and models of machine/deep learning systems is of distinguished importance nowadays. Among the first efforts in the fields of knowledge-intensive systems and artificial intelligence development and use, was the paper of Chandrasekaran (1986). This seminal work addressed the use of high-level structured knowledge blocks for expert systems. Attempting to move beyond the capabilities of contemporary KBSs mandates knowledge bases that are substantially larger than those we have today. McDermott (1990) described how artificial intelligence research could make software development easier by writing programs “to act as frameworks for handling instances of problem classes in software engineering”.

The need for and the possibility of knowledge exchange between engineered systems was also addressed in the seminal work of Neches et al. (1991). They identified three possible forms of knowledge sharing: (i) communication of the principles of knowledge bases to facilitate their reimplementation, (ii) facilitation through the inclusion of source specifications into new knowledge components, and (iii) run-time invocation of external modules or services. On the other hand, they also identified four impediments: (i) heterogeneous representations, (ii) dialects within language families, (iii) lack of communication conventions, and (iv) model mismatches at the knowledge level. Smith and Poulter (1993) recognized the need for open knowledge-based systems and proposed an open infrastructure that permitted the integration and interoperability of different **knowledge-based systems** (KBS) and ensured that each system could utilize whichever representation for knowledge is appropriate to its tasks. As elements of an open KBS infrastructure, they defined: (i) standard knowledge representations, (ii) knowledge interchange format, (iii) knowledge manipulation and query language, (iv) common shared ontology, and (v) agent-based software engineering framework. The open KBS infrastructure supported run-time sharing of complexly structured knowledge between knowledge bases and their associated inference engines even if they used different knowledge representation formalisms and different inference mechanisms.

There were parallel efforts that yielded the Initial Graphics Exchange Specification (IGES) and the now international standard (ISO 10303) Standard for Exchange Product Data a decade later. The latter has been under development since 1984 and in use since 1994 (Pratt, 2005). The initial parts of the standard were orientated towards transferring voluminous **artifact and process model data** (CAD CADE, CAPP, and CAXX data) between multiple design and engineering systems using neutral representation formats. The latter parts, such as the ISO 10303-239 (application protocol for product life cycle support - STEP PLCS), have covered the entire product development and use process from conceptual design to recycling.

Like other productive resources, system intellect resources can be (i) shared among similar systems, (ii) adapted and combined on purpose, (iii) warehoused and archived, and (iv) retailed as a cognitive product. Over the years, many technologies have emerged that can support the real-life implementation of all of these general options. At the same time, exchange and reuse of system intelligence of i*CPSs has not obtained due attention in the literature yet, nor has it been addressed in large-scale projects. In principle, it can happen: (i) in a human-assisted manner, (ii) in a systems-planned autonomous manner, and (iii) in a hybrid manner. Since there is a high probability of an autonomous extension of the functional profile of i*CPSs, SSI transfer may become a practical technological solution for obtaining the needed intellectual resources. However, it should be seen as a partial solution because the whole spectrum of resources (interoperating analogue and digital hardware, system-level and application-oriented software, and signals, data and information) are to be availed (acquired in run-time) too. Runtime resource management is the major issue for adaptive and, in particular, for evolving i*CPSs.

7.3. System intellect as a new industrial asset

The SSI self-acquired or self-generated by intellectualized engineered systems is becoming an important complement of human knowledge and problem-solving intellect. In addition, generating, exporting and/or importing, and reusing SSI via intellectualized systems have become possible technologically and necessary economically. Contrary to its growing importance and volume, the general methodological underpinning and the strategy of practical exploitation of SSI are still underdeveloped. Actually, the whole field of interests has not received sufficient attention. This situation triggers many questions to which the studied literature fails to give convincing answers.

According to the traditional interpretation of the knowledge transfer problem, there is a need to identify the **highest common denominator** among the knowledge representation and interchange mechanisms of the systems to be integrated. In contrast with this, the completed research pointed at the opportunity of applying a different approach to utilizing system intelligence, which may be based on the principle of 'share it if you have or need it'. For instance, recommender systems may collect information about the exchangeable SSI as well as meta-information about the demands and supplies. One issue of taking steps towards exchanging and using SSI as a system-independent asset is disconnecting it from the original producer system. This warrants attention because, on the one hand, SSI should be made transferable in the simplest way, and, on the other hand, its problem-solving power should be preserved as much as possible. It calls for intelligence-oriented networking of i*CPSs, well beyond their physical and communicative networking. This is a new **problematics**, which has yet been scarcely studied, likewise the issue of adaptation and operationalization of the knowledge and mechanisms constituents of SSI imported from multiple source systems to one target i*CPSs.

In general, there is a strong interconnection between the problem-solving knowledge and the processing mechanisms of i*CPSs. Typical examples are such as: (i) production rules <—> inference engine, (ii) analogy-based cases <—> case comparator, (iii) fuzzy rules <—> fuzzy reasoning engine, (iv) chromosome constructs <—> genetic algorithms, and (v) training data sets <—> artificial neural network. Due to their inseparable nature, these constituents of the SSI should be captured together and considered as duals for an effective SSI transfer. At the same time, this **duals-orientated intellect transfer** needs different packaging mechanisms than the neutral interchange format-based mechanisms typically applied in the case of traditional knowledge-based systems. It can be assumed that i*CPSs will be able to computationally manage such packaging mechanisms of duals.

This above-sketched inter-systems knowledge sharing may open up a new direction for utilization of SSI and may amplify the problem-solving potential of i*CPSs as has happened with human knowledge and companies. Some pioneering researchers believe that such cooperative systems, or systems of systems, have the potential to be a game changer in multiple creative and productive domains. For the purpose of this work, intelligence transfer is understood as all of the structured activities related to separating application-specific knowledge and processing mechanisms from one system and embedding them into several interoperating systems. Due to the obvious space limitations, the main features of the particular approaches can be presented only from a birds-eye-view. However, this is deemed sufficient to understand the overall conceptual relationships and the logic of reasoning.

7.4. Transfer based on repositories

The first examples of transferring intellect between systems are related to the symbolic and analogical methods of artificial intelligence research and system development. During the 1980s, it was recognized that building new knowledge-based systems usually entailed constructing new knowledge bases from scratch. Therefore, the scope of the built systems remained restricted, their development needed a lot of time, and the costs and efforts ran high. As a solution, proposals were made to assemble reusable knowledge components by system developers and to make succeeding systems able to interoperate with existing systems and use them to perform some of their reasoning tasks. In this way, declarative knowledge, problem-solving techniques, and reasoning services could all be shared among systems. However, both specific technologies and sophisticated infrastructures are needed to realize the repository concept on a large scale.

The idea of **repository-based knowledge exchange** has gone through a number of developmental stages, such as (i) database sharing, (ii) semantic networks, (iii) symbolic rule-fact bases, (iv) analogical example libraries, (v) relational knowledge-bases, (vi) resource description frameworks, (vii) web ontology languages, and (viii) knowledge fusion frameworks. Common characteristics of these milestone concepts are that they (i) formalize and structure human knowledge chunks, (ii) assume various description or specification languages, (iii) provide opportunity for external queries, and (iv) are not, or not directly, related to application cases. These approaches allow transferring knowledge content from the repository to one system, but do not support direct knowledge sharing among application-oriented systems. Not only the knowledge engineering process, but also the knowledge acquisition (retrieval or extraction) process is human dependent.

The pioneering knowledge transfer approaches rested on agreements concerning (i) a standard syntax and semantics, (ii) a knowledge interchange format, (iii) a set of protocols to query a virtual knowledge base, (iv) a common shared ontology content, (v) a vocabulary and constraints on the well-formed use of contents, and (vi) an agent-based software engineering framework. Typical implementations concern (i) computational routine libraries, (ii) chunks of procedural knowledge, (iii) rule interchange format, (iv) labelled case libraries, (v) annotated object repositories, and (vi) product and service catalogues. The examples indicate that repository-based transfer may include both content knowledge and processing mechanisms. In practice, repository-based SSI transfer may concern three purposes (i) transferring SSK only, (ii) transferring reasoning mechanisms only, and (iii) transferring synthetic systems intellect. Artificial intelligence research comprehensively studies the various application-independent, but task-driven forms of computational reasoning approaches.

7.5. Transfer among agents

The second example is taken from the field of **multi-agent collaborative systems**. Multi-agent systems are decentralized structures formed by autonomous computational entities that communicate and share data, information, and state-maintaining and problem-solving knowledge with each other (Leitão et al., 2016). Agents represent real-world or virtual entities with varying levels of fidelity, intellectualization, commitment, and socialization. They are implemented as intellectualized entities, which have sufficient intellect and capacity for (i) building situational awareness, (ii) making logical decisions, and (iii) performing functional agency. Therefore, many works consider them ‘intelligent’ entities (Rudowsky, 2004). Informally, their rational intelligence is seen as the ability to achieve goals in a complex environment, whereas their social intelligence is the ability to successfully interact in an environment full of other agents (Insa-Cabrera and Hernández-Orallo, 2013). Agents act, learn, negotiate, and adapt autonomously and try to understand their environment in order to pursue their goal. With regard to the autonomy of the agents, important issues are goal delegation and goal adoption, which are seen as ingredients of organization, social commitment, and contract of the agents, and then of the knowledge exchange process.

In multi-agent systems or system-of-systems, the issue of transferring intelligence, resources, and/or knowledge from one actor agent to others has been known for a long time (Sycara et al., 1996). To communicate, the agents are supposed to comply with output guaranties and input assumptions - otherwise their interoperability cannot be provided. While there are papers discussing in-process communication and information exchange among agents, much less is published on transferring aggregated knowledge from one agent to others (Allen et al., 2002). The inter-agent transfer of intelligence among hardware and software agents concerns not only signals, data, and pieces of information, but also chunks of knowledge and experiences that they have individually learnt according to their operation strategy. Transfer can be initiated both by a receiver agent and by a sender agent when knowledge and logic required to perform an activity is unavailable, incomplete, or out of date. The knowledge shared by the agents constitutes beliefs proven individually or collectively by the collaborating agents. As a result of this, multiple collaborating agents build distributed intelligence.

The transfer of the necessary knowledge and processing logic may enable an agent to execute a new task or to execute a given task better (Iglesias et al., 1998). The process includes three main steps: (i) packing and decoupling the knowledge and algorithms from the sender agent, (ii) routing and transferring, and (iii) unpacking and embedding in the receiving agents. Alternatively, when the knowledge is processed by local algorithms of the receiver systems, (iv) activation of the various local algorithms has to happen. As a whole, the multi-agent system may harmonize the package sending and processing over all sender and receiver agents, or may leave it on their own decision which depends on their programmed objectives, social character, and local context (Cardoso and Ferrando, 2021). As explained by O'Neill and Soh (2022), the subsequent steps of the process are: (i) triggering the messaging actions either by a pre-programmed timer poll or by an event-drive framework, (ii) building local situational awareness according to data obtained from own sensors or memory, or received by communicating with other agents, (iii) understanding the meaning of between-agent communications (Williams, 2004), (iv) making a decision based on the logical image and the built situational awareness, (v) execution of the decision locally or in cooperation with other concerned agents, (vi) consulting as an action carried out locally or potentially by some other agent through a cooperative or delegated process. The level of situation awareness depends on whether the agent only aggregates data, or actuates its functional model by time-wise obtained data.

Smart agents can migrate from one system to another, taking their knowledge with them, and can continue their operation after hospitalization in the target system from where they left off. The principle of **agent hopping** as a transfer mechanism may contribute to finding a solution to the system intelligence utilization problem where the system's actors can be agentized. The pioneers, such as Bharat and Cardelli (1995) developed the principles of how application migration could be implemented at the programming language/environment level. They proposed to include two complementing elements, namely suitcases and briefings. A suitcase is the long-term memory of the agent that contains all pieces of knowledge that the agent can take with it. It may include own-knowledge to share and own-tasks to execute. The briefings are chunks of knowledge that the migrating/migrated agent receives from the target system. The contents of these containers are updated before every migration. Thus, suitcases and briefings are the enablers of the interoperation of an agent with other hosting agents (Xu and Qi, 2008). Smart software agents of i*CPSs can diagnose the opportunities for migrating and can make decisions on the execution and timing of a migration autonomously, based on the possessed data and obtained communications (hop instruction). Agents may duplicate themselves and send their copies to multiple target systems. A recognized issue is that agents with diverse ontologies may assign different meanings to the same concept, or consider different concepts and messages with the same meaning (Athanasiadis, 2005). Coordination of nearly concurrent migrations of agents is an additional computational issue, as well as the negotiation protocol development for autonomous multi-agent systems.

7.6. Transfer of learning resources

A third evolving example of technological opportunities for knowledge transfer between intellectualized systems is the transfer of learning resources (data, models, mechanisms, rules) (Zhuang et al., 2020). Actually, two complementary forms of computational learning deserve attention. One approach, referred to as **transfer learning**, is associated with the recently developed sophisticated computational mechanisms of machine and deep learning (Neyshabur et al., 2020). The most basic form of transfer learning is fine tuning a pre-trained model. In addition to the mentioned transfer learning, federated learning also deserves attention. It also aims at transferring learning resources, but differently (AbdulRahman et al., 2020).

Transfer learning has been proposed to enable modification of the architecture of deep learning networks by pre-trained layers and to solve the problem of insufficient training data (Pan and Yang, 2010). The architecture modification is typically a focused action, i.e. it involves editing only the last few layers of the target network, as opposed to modifying the layers in the command line. This way, the output functionality of a network can be changed, for instance, from classification to clustering, as happens in the case of MatLab (Figure 7.2). This form of computational learning also supports cases where training data are expensive or difficult to collect, and makes it possible to provide labelled deep learning data or to change the algorithmic elements of the learning mechanism (Torrey and Shavlik, 2010). In this way, data and algorithms can be used more efficiently and aptly in other application cases, having a number of characteristics in common and being not prone to negative transfer. As an example of the latter, Zhang et al. (2018) posited that recommender systems often suffer from the data sparsity problem that is prevalent in newly-launched systems having had not enough time yet to amass sufficient data. To solve this knowledge insufficiency problem, these systems apply cross-domain knowledge transfer (i.e., transfer relevant data and relationships from a rich source domain to assist recommendations in the target domain).

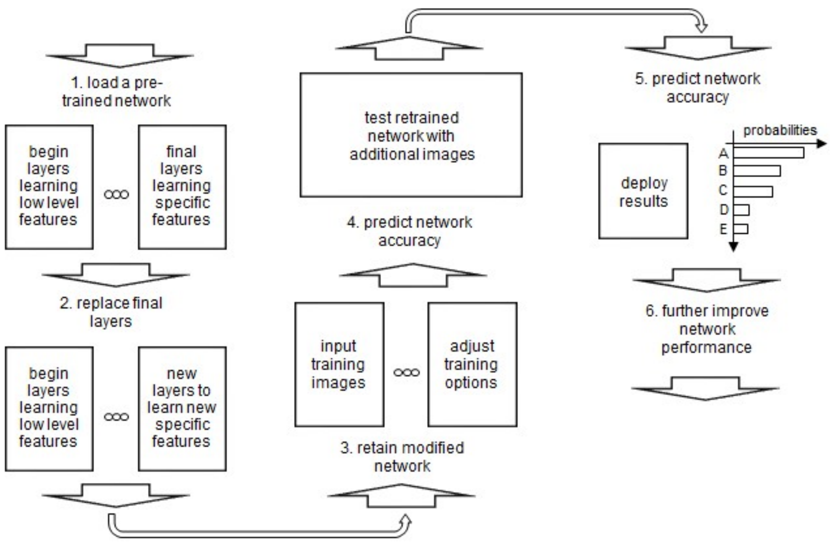

Figure 7.2.: The overall procedure of reusing pre-trained layer components of deep neural networks

Like **multi-task learning**, transfer learning also exploits relations between different learning tasks (Zamir et al., 2018). In contrast to multi-task learning, which simultaneously (jointly) solves many related individual learning tasks, the methods of transfer learning operate in a sequential fashion and solve the learning tasks consecutively. Transfer learning is enabled by constructing regularization terms for a learning task by (re)using the results of a previous learning task (Weiss et al., 2016). A popular implementation is deep transfer learning. Deep learning mechanisms attempt to learn high-level features from mass data by automatically extracting data features by unsupervised or semi-supervised feature learning and hierarchical feature extraction, and to use the learnt features to classify objects. Deep transfer learning is often categorized based on the computational approaches used. Based on these, the following categories are identified: (i) instances-based (utilizing instances in the source domain by appropriate weight), (ii) mapping-based (mapping instances from two domains into a new data space with better similarity), (iii) network-based (reuse the partial of network pre-trained in the source domain), and (iv) adversarial-based (use adversarial technology to find transferable features that both suitable for two domains). Deep learning has a very strong dependence on massive training data compared to traditional machine learning methods.

The term '**federated learning**' (FL) was first introduced in McMahan (2016) to name the approach of collaboratively training a machine learning model based on distributed resources. Strictly speaking, FL is an umbrella term for ML/DL methods that train models in a collaborative fashion. Opposing the other centralized approaches, FL is a distributed machine learning approach, which keeps raw data decentralized, or in other words, without being moved to a single server or data center (Khan et al., 2021). That is, the mechanism of FL brings the code to the data, instead of bringing the data to the code. On the other hand, it coordinates the distributed trainers to carry out the training process of machine learning efficiently (Sattler et al., 2020). There are three aspects in which FL differs from other centralized learning approaches: (i) it allows transferring the learnt (intermediate) data among the distributed computing resources, while avoiding the transfer of training (direct raw) data, (ii) it exploits the distributed computing resources in multiple regions or organizations, (while the centralized approach generally only utilizes a single server or a cluster in a single region, which belongs to a single organization), and (iii) FL generally takes advantage of encryption or other defense techniques to ensure the data privacy or security, while the centralized approach pays little attention to these security issues (Smith et al., 2017).

FL is formally defined as a machine learning approach where multiple clients collaborate in solving a **machine learning problem**, while the raw data is stored locally and is neither exchanged nor transferred (Shaheen et al., 2018). Federated learning does not allow communication (while the centralized approaches have no restrictions), because it addresses the fundamental problems of privacy, ownership, and locality of data. The concept of FL was extended to three data scenarios, i.e., horizontal, vertical, and transfer (Zhang et al., 2021). The distributed machine learning implemented according to the horizontal data scenario of FL addresses decentralized data of the same features, while the identifications are different. Features are those properties (predictors) of a data construct that can be measured or computed in an automated fashion. For example, colors are features of a pixel in a bitmap image. The vertical data scenario handles decentralized data of the same identifiers with different features. The hybrid data scenario deals with data of different identifiers and different features. Network coding techniques have been applied to the design and analysis of FL methods (Sarcheshmehpour et al, 2021). The various approaches of transfer and federated learning offer mechanisms that can be used as analogical in the case of system intelligence and knowledge transfer.

7.7. Transfer by emerging approaches

As a fourth example of the current approaches of knowledge transfer between intellectualized systems, computational approaches of knowledge graphs (KGs) reuse (Hogan et al., 2021) and collective intelligence have been taken into consideration. The term ’**knowledge graph**’ (like the term ‘semantic network’) was introduced in the literature at the beginning of 1970s (Schneider, 1973). It has been revitalized by commercial companies at the beginning of 2010s (Noy et al., 2019). The reason for this revival is that graphs provide an intuitive and concise abstraction for a variety of knowledge domains, where nodes, edges, and paths capture different, potentially complex relations between the chunks of knowledge. Knowledge bases and knowledge graphs show some similarities. The relational records of a knowledge base are replaced by single- or double-orientated entity-relation/predicate-entity constructs of knowledge graphs. Though versatile, KGs are also not always sufficient for problem-solving by i*CPSs (Abu-Salih, 2021). That is why they have many different forms of extension mechanisms.

Conceptually analogous to a **non-hierarchical concept map**, KGs are seen as more complex than image or text data types, which are characterized by (i) lack of reference points, (ii) arbitrary size, and (iii) diverse network topology. In principle, knowledge graphs can be constructed without any underpinning predefined ontology schema. Nevertheless, the literature does not report on computational methods for automated graph construction, only on graph processing (e.g., on transformation of entities and relations into a continuous vector space). In this arrangement, the knowledge graph is the knowledge container and the machine learning mechanisms may avail its content for use by different systems. Typical models of KGs are (i) directed edge-labelled graphs, (ii) heterogeneous information networks, (iii) entity-property-value graphs, (iv) graph meta-datasets, and (v) stratified hyper-graphs. Reasoning can be (i) inductive symbolic reasoning (e.g., self-supervised rule-mining and axiom-mining), and (ii) inductive numeric reasoning (e.g., unsupervised graph analytics, self-supervised embeddings, and supervised graph neural networks).

Since the real-world knowledge graphs are large and highly incomplete, inferring new facts based on them is challenging. Being a network of entities and their relations in their simplest form, KGs embed discrete but linkable elements of knowledge and can be extended with various reasoning and learning mechanisms (Tiwari et al, 2021). Direct processing of knowledge graphs includes (i) knowledge graph embedding in vector spaces, (ii) knowledge representation learning, (iii) knowledge graph completion, (iv) extraction of relation paths, and (v) knowledge graph completion (Ji et al., 2021). The knowledge graphs stored on a cloud, a fog, or an edge are actually not shared, but directly accessed by multiple systems even concurrently.

Machine learning mechanisms can learn the interrelated knowledge hiding in the relational structures within domain-specific or domain-independent heterogeneous KGs (Tian et al., 2022). By embedding a graph in a vector space, its logical representation can be transferred to (a dense) numerical representation. For instance, Liu et al. (2022) extended a given knowledge graph representation with machine learning. The encoding of KGs can be executed by deep learning through relational **graph convolutional network** (GCN). The entity and relationship constructs are embedded using translational models such as TransE, ConvE, ComplEx, RotatE, QuatE, and AutoSF. Many researchers share the opinion that knowledge graphs can become a confluence of technologies from different areas with the common objective of maximization of knowledge which can be distilled from diverse sources at a large scale using a graph-based data abstraction (Hur et al., 2021)

As discussed by (Lykourentzou et al., 2011), the idea of **collective intelligence** (CI) and **collective intelligence systems** (CISs) has emerged in the context of producing higher-order intelligence, solutions, and innovation by large groups of cooperating individuals. However, the attention has twisted to the implementation of synthetic collective intelligence in the last decade. In the formulation of Sulis (1997), a CIS consists of a large number of quasi-independent, stochastic agents, interacting locally both among them, as well as with an active environment, in the absence of hierarchical organization but in the presence of adaptive behavior. The three principles (stochastic determinism, interactive determinism, and nonrepresentational contextual determinism) and the two major behavioral control processes (non-directed communication and stigmergy) he identified in a different context, have logical links to swarms of systems and their swarm intelligence. Gunasekaran et al. (2015) proposed a theory of collective intelligence that mimics the communication process typically occurring in the collaboration of human entities in self-managing multi-actor systems. It attempts to explain the emergence of intelligent collective behaviors, among others, in social systems. Musil et al., (2015) proposed a multi-layer model that includes three constituents: (i) human actors as proactive components, (ii) a single, homogeneous CI artifact network as a passive component, and (iii) reactive/adaptive component for computational analysis, management and dissemination.

Passive and active CISs have been distinguished. Zhang and Mei (2020) presented a constructive model for collective intelligence, which continuously executes exploration, integration, and feedback in computational loops. The idea of CISs can be extrapolated to the transfer of synthetic system intelligence based on adaptation of the previously proposed approaches and introducing new ones. Artificial collective intelligence is seen as a new perspective on AI, which is enriching computational intelligence techniques (Williams, 2021). The latest implementations of this technology seek to merge human and machine intelligence with the aim of achieving results unattainable by either one of these entities alone (Smirnov et al., 2019). From a practical point of view, it facilitates achieving the goals of a multi-actor system at a collective (group or crowd) level. The elements of the overall knowledge transfer process are (i) discussion, (ii) argumentation, (iii) negotiation, and (iv) decision making.

Leitão et al. (2022) completed an extensive literature study concerning the concept and features of collective intelligence in an agent-based CPS. According to them, the concept of collective intelligence provides an alternative way to design complex systems with several benefits, such as modularity, flexibility, robustness, and re-configurability to condition changes, but it also presents several challenges to be managed (e.g., non-linearity, self-organization, and myopia). What differentiates CISs from multi-agent systems is that the shared knowledge is transformed, cross-fertilized, moderated, and consolidated through a series of discursive interactions among the actors. Notwithstanding, each included entity has its own **personal intelligence**. Chunks of crowdsourced information and collective intelligence can be used as input into learning mechanisms. Zheng et al., (2018) proposed a computational platform to support the development of multi-agent-based reinforcement learning for artificial collective intelligence.

7.8. Technological framework for managing SSI

Sections 7.3 - 7.7 presented seven already consolidated or currently developing computational technologies that show affinity to a comprehensive and robust SSI transfer process as well as to each other. These are: (i) distributed (intelligence) repository management, (ii) collaborative multi-agent-based transfer, (iii) migrating multi-agent-based transfer, (iv) transfer of learning resources, (v) decentralized federated learning, (vi) knowledge graphs-based transfer, and (vii) collective intelligence-based transfer. It is probable that none of them alone will be sufficient for all SSI transfer problems of i*CPSs. As an intermittent stage in research and development, a subset, the whole set, or an extended set of the above-discussed technologies should be integrated procedurally and computationally into tailored and orchestrated **transfer technologies** (Figure 7.3). It should also be taken into account that there are several - yet in sprouting - technologies, such as knowledge transfusion and knowledge distillation, but

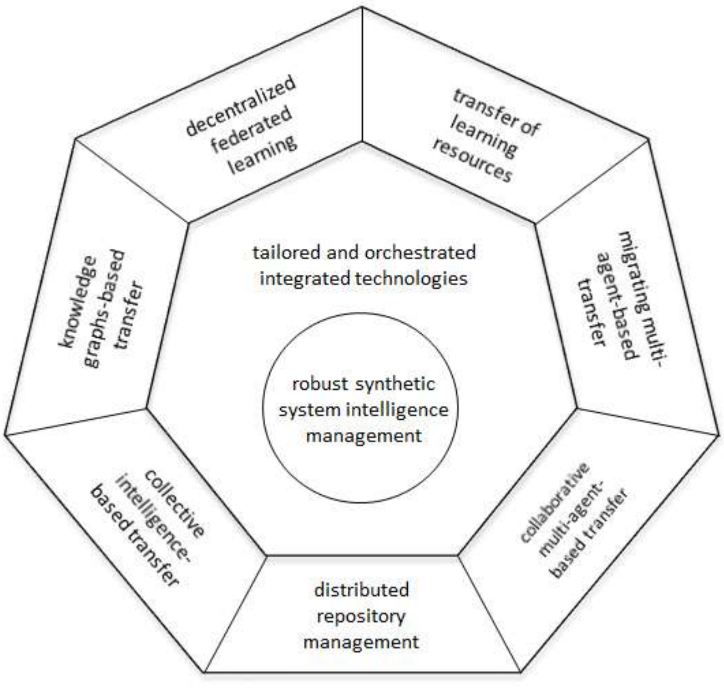

Figure 7.3: Potential technological processes of utilizing SSI as an industrial asset

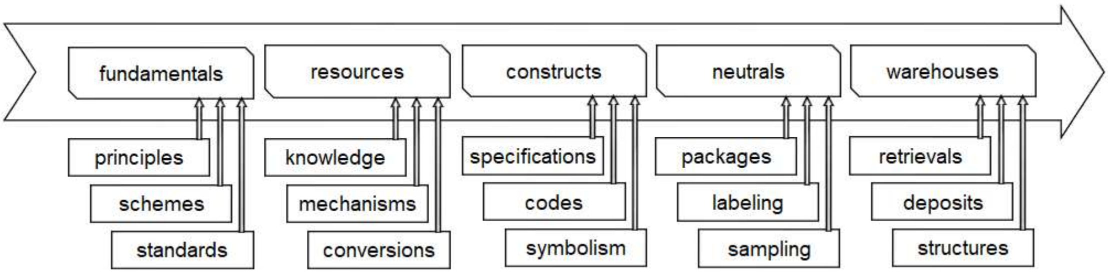

Figure 7.4: The technology related conceptualization of utilizing SSI as industrial asset

it is not clear how they can contribute to solving the problem of transferring synthetic system intelligence.

It is fair to mention that this set of probable technologies reflects a (subjective) conjecture of the author. Nevertheless, these technologies are supposed to get further developed by domain experts and to reach that high level of maturity, which is required in

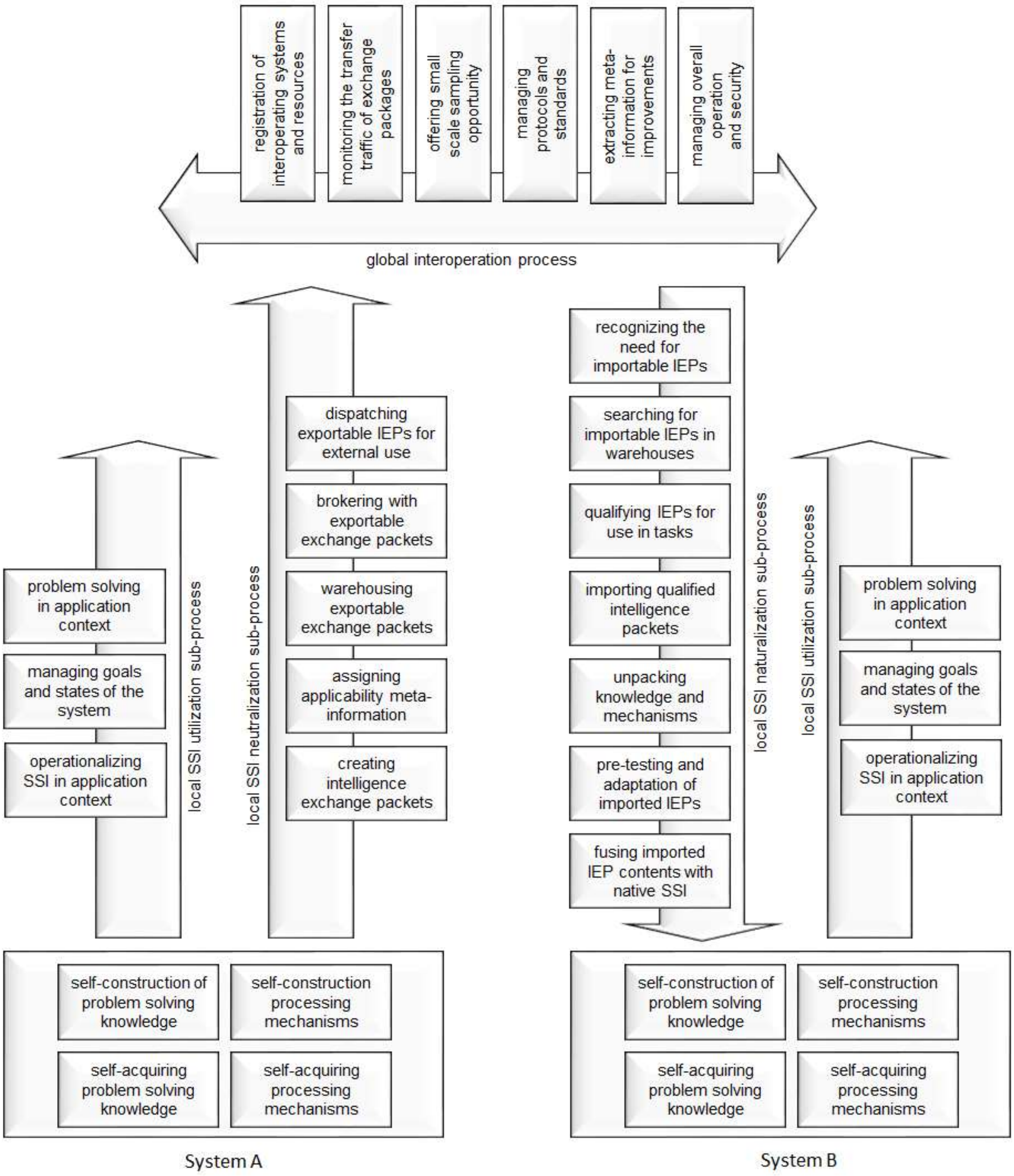

Figure 7.5: The technological processes of utilizing SSI as industrial asset

the context of transferring SSI. The main concepts underpinning the utilization of SSI as an industrial asset are shown in Figure 7.4.

From a technological point of view, the overall **asset management process** of SSI can be divided into local processes and a global interoperation process (Figure 7.5). The outcome of the global process is the generic SSI distributed over all interoperating systems. There are two types of local processes named **neutralization process** and **naturalization process**. The neutralization sub-process is about separation of transferable SSI from the generating system, while the naturalization sub-process is about the integration of transferred SSI with the native SSI of a system. The outcome of the local neutralization sub-process is a package of SSI self-generated by a source system (called **exported SSI**), whereas the outcome of the local naturalization sub-process is an SSI packet integrated with the native SSI of the target system (called **imported SSI**). Both the export and import SSIs include task-oriented problem solving knowledge and processing mechanism combinations.

The neutralization sub-process is a (tail) extension of the local intelligence management process. Thus, this extends: (i) self-acquiring problem solving knowledge, (ii) self-acquiring processing mechanisms, (iii) self-construction of problem solving knowledge, (iv) self-construction processing mechanisms, and (v) operationalizing SSI in application context with (vi) creating **intelligence exchange packets** (IEPs), (vii) assigning applicability meta-information to IEPs, (viii) warehousing exportable IEPs, (ix) brokering with exportable IEPs, and (x) dispatching exportable IEPs for external use.

The naturalization sub-process is a (front) extension of local intelligence management processes. It appends (i) recognizing the need for importable IEPs, (ii) searching for importable IEPs in warehouses, (iii) qualifying IEPs for use in tasks, (iv) importing qualified IEPs, (v) pre-testing and adaptation of imported IEPs, (vi) integrating the contents of imported IEPs with native SSI used for problem solving activities and self-management activities by the host system. The activities of the global transfer sub-process are: (i) registration of interoperating systems and their resources, (ii) handshaking and monitoring the transfer traffic of IEPs, (iii) offering small scale sampling opportunity, (iv) managing protocols and standards, (v) extracting meta-information for improvements, and (vi) managing overall security.

Implementation of SSI transfer means extra overheads for intellectualized systems from four aspects: (i) operationalization, including (a) pretesting, (b) integration, and (c) refinement, (ii) long-term wrangling, including (a) evaluation, (b) filtering, (c) chunking, and (d) extension, and (e) structuring, (iii) enrichment, including (a) annotating, (b) contextualization, and (c) tailoring, and (iv) packaging, including (i) assembling, (ii) labelling, and (iii) standardization.

7.9. Provisioning as an industrial asset

The increase in industrial revenues and social benefits poses a continual need for novel innovations and new assets. Traditionally, an asset is a resource owned and controlled by an individual, a production or servicing company, or a government. It is a result of past or current activities, the enabler of economic benefits. In the past, multiple forms of human knowledge (scientific, technological, enterprise, educational, etc.) have been used as industrial assets. What constitutes human knowledge assets are (i) the outputs of the knowledge transformation processes, and (ii) the accumulated depository of skills, knowledge and experience of human professionals. Knowledge produced by artificial intelligence has also reached this status.

In comparison with the conventional assets, SSI has unique characteristics since it is: (i) intangible, (ii) sharable, (iii) reproductive, (iv) evolutionary, and (v) context-valued. It can be possessed as a property, and/or accessed as a service. Thus, SSI contrasts the traditional (narrow sense) interpretation of industrial assets as means (equipment, tools, chemicals, vehicles, infrastructure, computers, materials, etc.) deployed to convert inputs to industry outputs, which can then be marketed as products, services, and experiences, with the expectation that they will generate future cash flows. Handling SSI may become a part of the practice of asset management, because it has the potential to grow in volume and value, and to increase total wealth over time. The reasoning regarding the logic of provisioning SSI as an industrial asset may start out from the key properties of an asset. Typically, three properties are identified: (i) ownership/access, (ii) economics, and (iii) supply. In addition to the technological and business issues, these important issues also need further attention.

For instance, **ownership** seems to be a simple matter in view to the current status quo of engineered systems and the concerned legal regulations, but in fact it is not. According to the latter, the responsible owner of the SSI is the original developer and/or the actual owner of the system producing the asset, as contracted. However, this is not so straightforward in the case of autonomous systems of the near future, which produce their synthetic resources/assets largely independent of humans, or at least not under the direct control of human stakeholders. The other side of the coin of the ownership of SSI is that (i) proprietary, (ii) shared, and (iii) open forms of possession may take place. Proprietary SSI means that the body of knowledge and the processing mechanisms belong inseparably to a system (or to the owner of similar systems). This knowledge is primarily stored in the repositories of the system or on those of the owner company (e.g. on edge computing devices or a private cloud of the company or third party, with no or limited access to other enterprise and partner networks, and retaining a high degree of control, privacy, and security). **Shared SSI** means that the body of knowledge is jointly aggregated by cooperating systems and/or their owner companies over multiple systems. It is managed either on shared edge networks or on a community cloud whose infrastructural elements and processing rights are shared by several organizations or third parties which share concerns, common objectives, and optimization of benefits. Open SSK means that a body of knowledge is made openly accessible, processable, and usable for the systems, developers, and researchers of a large industry group, academic organization, and eventually, the broad public. In other words, the historically aggregated and maintained intelligence may reside on publicly accessible clouds or may be availed by a cloud service provider, enabling standardized data and application portability.

As well, the economic value of the exchanged, sold, or obtained SSI assets is a complicated matter (Amin et al., 2018). Assets are associated with ownership and can eventually be turned into cash and cash equivalents for the owners. Ultimately, it means that the total amount of investments should be less than the total amount of financial return (profit). The investments include all (primary) costs of (i) the implementation of the system shell, (ii) the knowledge engineering in the set-up stage, including the preparation of the reasoning and control mechanisms, (iii) the processing (extraction) of system intelligence during operation of a source system, (iv) transferring system intelligence to a target system or to a warehouse, and (v) reactivation of the transferred system intelligence in a target system. The returns include the (primary) income based on (i) selling and maintaining the system shell, (ii) vending knowledge engineering means and services, (iii) selling system intelligence, (iv) sharing the benefits of reusing transferred system intelligence by the target system(s). Here, only the direct and indirect costs and benefits are thought of, and the secondary costs and benefits are ignored. Both the investment side and the return side involve complex activity flows, whose financial consequences are difficult to capture in detail and, therefore, comprehensively forecast and quantify. The evaluation is even more complicated if multiple (large number of) systems (or a dynamic system of systems) are considered which have different commitments and involvement in asset generation and utilization, and may show different levels of successful and unsuccessful operations.

The supply aspect of converting SSI into an industrial asset involves not only opportunities, but also challenges. Traditionally, the concept of asset convertibility is used to classify assets according to how easy or difficult they are to be supplied and to get converted into cash. The primary issue would be the motivation of the owners of autonomous systems to make positive decisions on collaboration and to equip their systems up-front, or augment them in use time, with facilities for SSI management. However, utilization of SSI as a novel industrial asset is supposed to happen not only over the boundaries of systems, but also over the borders of companies and enterprises. This novel form of asset exploitation is deemed to be part of their information technological (IT) asset management. It must complement the **combined practices** of technological, financial, inventory, and contractual functions within the IT environment, and help strategic decision-making, optimization of spending, and support lifecycle management.

7.10. Reflections and open issues

In a sense, history repeats itself: In the 1970s-1980s, the need for technological solutions for transferring product data and inference knowledge among dissimilar systems was a stimulant of information systems research. In the years of the 2010s and 2020s, the need (and the opportunity) for technological solutions to transfer synthetic knowledge among intellectualized systems has given orientation to a part of CPS research. The overall assumption of the background research was that shared SSK and reasoning/learning mechanisms can eventually become a valuable industrial asset. As an intellectual capital, SSI can contribute to the net working capital of a company or even to the problem-solving potential of the whole society. This chapter of the disseration was intended to present the thoughts of the author concerning a number of recognized issues and technological affordances – with an obvious incompleteness.

It is well-known to scientists that the process of learning and knowing a yet unknown research phenomenon goes through such stages as discovery, description, explanation, prediction, and regulation. Since the emerging phenomenon of managing synthetic system intelligence is novel even on a conceptual level, the research could focus only on the identification and the characterization of this complex, multi-faceted phenomenon. Therefore, a larger part of the contents presented in this position paper belongs to the stage of discovery and description, and only a smaller part to the stage of explanation and prediction. Due to the newness of the phenomenon, the literature is rather scarce and unspecific. Thus, the intention of the structured literature study was to get deeper insights into the state of science and practice.

- The completed survey informed us about the lack of generic theories, conceptual frameworks, and methodological approaches. Many of my own concepts and ideas are only work-in-progress and they still must be addressed extensively from the perspective of an industry-wide implementation. Thus, a supplementary goal of the work has been to stimulate and encourage research and development efforts in this direction. As good research questions imply new questions, cutting-edge prognostic research should encourage many strands of follow-up research.

- Like patents and copyrights, SSI is to be treated as a (i) partially-physical, (ii) intangible, (iii) liquid, (iv) functional, and (v) net identifiable potential asset that needs a socio-technical process to get converted into cash. The road to an industrial solution for utilization of SSI will most probably be long, curvy, and bumpy. Nevertheless, it is wise to deal with it in the framework of digital transformation, which has rapidly turned itself to a digital disruption in terms of intellectualization of engineered systems (Vial, 2021).

- This chapter has made an attempt to provide procedural framing of the process of transforming SSK into a common industrial asset and capital. Obviously, there are many fundamental unanswered questions concerning the vision of SSI and utilizing it as a new industrial asset. For instance: What is the true future of system intellect (or intelligence)? Where does SSK go? Where do the computational reasoning mechanisms go? But nothing can be an obstacle to imagination, design, and planning.
- Not surprisingly, this chapter also closes with questions, rather than with definitive answers. The main questions are: 'Have the investigated transfer mechanisms enough potential to serve a wide range of intelligence transfer problems of i*CPSs equally well?' Or, 'Should they be extended or substituted by some other (currently known or unknown) mechanisms that will serve better for the purpose?' However, these questions cannot be answered without implementation as a computational prototype.
- As the analysis showed, there are at least seven 'carriages' needed to move ahead on the road: (i) obtaining deep scientific insights into the overall phenomenon, (ii) elaboration of the fundamentals (underpinning knowledge and specifications), (iii) creation of conceptual, procedural, and methodological frameworks and models, (iv) working out the across-systems intellectual and computational mechanisms and resources of SSI transfer, (v) implementation of the technological, engineering, and organizational enablers, and (vi) realization of demonstrative prototypes with the involvement of autonomous i*CPSs, (vii) identification and propagation of the best practices among i*CPSs developers.

Chapter 8

8. Framing supradisciplinary scientific research for intellectualized cyber-physical systems

8.1. Diversification of cyber-physical systems

As touched upon in Chapters 2 and 4, the last decades have witnessed, and the forthcoming decades will most probably witness, the intensifying processes of scientific **convergence**, technology **integration**, and large scale **systematization** (Tanik et al., 2021). Offering new opportunities and affordances, the progression in fusing bits, atoms, neurons, genes, and memes in engineered systems is everyday evidence of this happening. It is conceived that this will lead to a **postdisciplinary** (holistic) science of systems and technologies. Likewise, the many current attempts to implement sophisticated CPSs for industrial and non-industrial applications are the indicators of the proliferation and implications of such a science. One of the induced phenomena is complexification, which leads to higher system complexities. The term 'system complexity' is used to express the total of a system's quantitative and qualitative characteristics, which can be traced back to the growth of (i) the elements of the systems, (ii) the relations among the elements, and (iii) the interactions of the elements (and/or the system as a whole) with the embedding local world. The new discipline of complexity science is emerging and addresses objectives such as: (i) getting cross-disciplinary insights into complex systems, (ii) explaining emergent structures and self-organization, (iii) generating effective abstractions and models, and (iv) providing control methods for complex systems. Current knowledge offered by complexity science is still in its infancy and unable to explain how to conceive, manage, and reduce aggregative complexities.

System complexity has two major constituents, which are physical complexity and cybernetics complexity. I have differentiated five types of **physical complexities**: (i) static complexity (the number and relationships of the elements that do not change with time), (ii) dynamic complexity (the number and relationships of components that change with time), (iii) self-organizing complexity (the variation of instances into which open systems can reorganize themselves), (iv) evolving complexity (the stage variations through which open systems can evolve in time into different systems) and (v) co-evolving complexity (two-way interplay between the changing system and its environment). These together are also referred to as aggregative physical complexity. **Cybernetic complexity** has three constituents: (i) the amount and kind of cyberware (data, information, knowledge, and experience) that is needed to design, implement, and sustain a system, (ii) the amount and kind of cyberware that is human-embedded into or self-acquired/generated by a system, and (iii) the amount and kind of cyberware a system is interconnected with the local world. The first item is related to the pluridisciplinary intelligence that is conveyed by all concerned scientific domains and is also referred to as disciplinary complexity. The second item is associated with the intellectual complexity of systems, which influences their smartness, autonomy, and evolvability of the system. Involving various modalities, channels, and levels, the third item points at a communicative complexity.

While the physical, virtual, and cognitive resources are converging, three trends of diversification of CPSs can also be recognized. They are termed: (i) intellectual diversification, (ii) disciplinary diversification, and (iii) application diversification (Figure 8.1). Intellectual diversification is happening due to the perpetual increase of cognitive

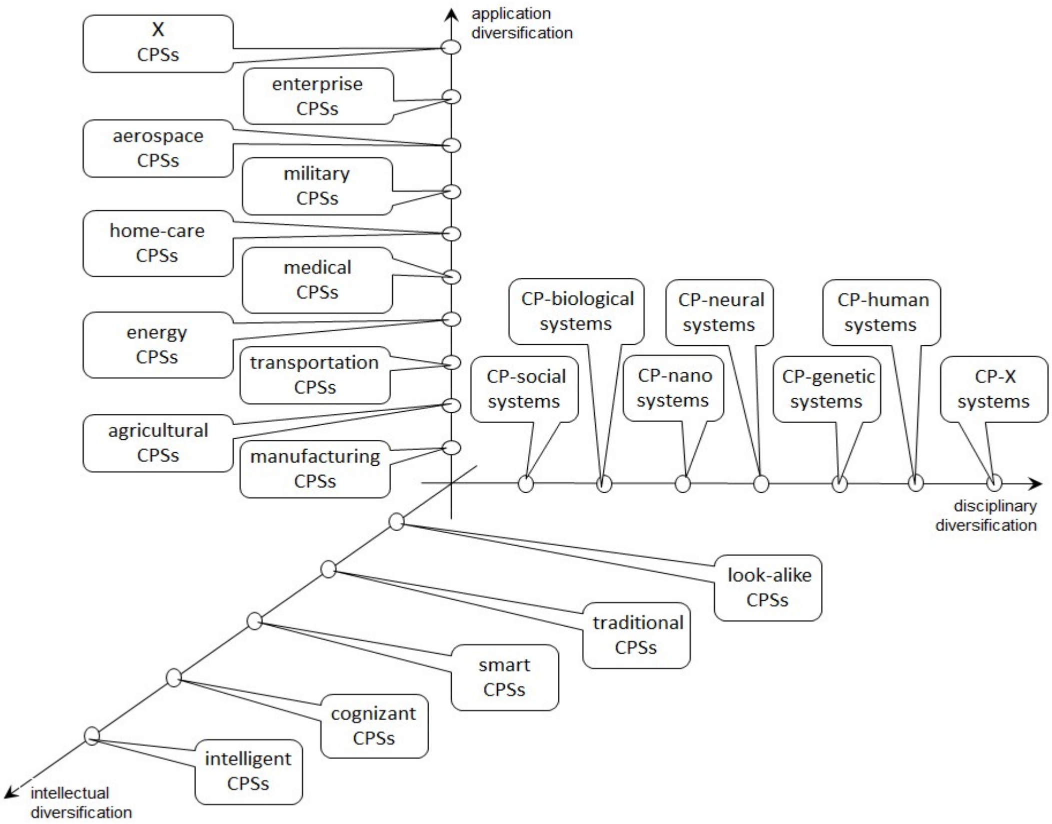

Figure 8.1: Dimensions of diversification of CPSs

capabilities of systems (Horváth et al., 2017). Current 2G-CPSs offer smart behavior based on their dynamic situation assessing and self-adaptation capabilities, whereas third generation CPSs will be characterized by some level of cognizance of the probable and possible objectives and performance, and will enhance themselves by self-evolution. Application diversification simply means that CPSs have a growing range of functionalities and, based on these, they are able to provide dedicated functional services for more and more application domains and penetrate into real life processes in these domains.

Strongly application orientated CPSs are typically distinguished by the acronym X-CPSs where X stands for the name (or only the first letter of the name) of the application field. Disciplinary diversification is associated with the variety of foundational knowledge and disciplinary resources that are used for the implementation of various CPSs. As mentioned above, bodies of knowledge of social, human, biological, nano, etc. disciplinary domains are considered in their conceptualization and design to make them suitable to tackle problems that could not be handled otherwise. Systems embodying such disciplinary bodies of knowledge are referred to as CPX systems, where X stands for the name of the discipline, for instance, cyber-physical-social and cyber-physical-cognitive systems.

8.2. Research objectives and approach

In the context of conceptualization and design of new CPSs, many socially-based problematics/phenomena have been emerging. Representative examples are such as: (i) coexistence of the natural intelligence-based and the artificial general intelligence-based realms of the society, (ii) addressing global societal demands in the light of the current technological affordances and economic situation, (iii) investigation of the role of i*CPSs in the fabric of Industry 4.0 and 5.0, (iv) exploring application opportunities for a family of smart (aware and adaptive) CPSs, (v) ideation of next generation home-care servicing CPSs with alternative cost profiles, (vi) aggregation and distribution of SSK from and to a fleet of

CPSs, (vii) intellect resource warehouses and upgrade mechanisms for self-evolving CPSs, (viii) complex solutions for transferable cyber-physical-social-human systems, (ix) effect of smart-everything systems of human life and well-being, and (x) opportunities for alternative living environments. All related in one way or other to cyber-physical-social-human systems, the textual formulation of these problematics suggests the need for different research approaches than those typical in system engineering practice nowadays.

The above changes have been creating new situations and challenges for both theoretical and empirical research. Consequently, CPSs research is currently one of the most flourishing fields of academic and industrial interests. Research in conceptualization and design of NG-CPSs is concerned with (i) integration of existing knowledge across the involved disciplines, and (ii) exploration and synthesis of novel knowledge by the involvement of all concerned disciplines. Based on literature studies, outcomes of workshop sessions, and critical mind mapping, it has been postulated that the most influential current trends of evolution of the paradigm of CPSs are: such as: (i) disciplinary complexification, (ii) synthetic intellectualization, (iii) behavioral socialization, and (iv) adaptive personalization (Figure 8.2). These pave the way to intellectualized, socialized, and personalized cyber-physical systems (cisp*CPSs). Such systems can be realized only by blending the relevant bodies of scientific, engineering, social, human, cultural, and application knowledge into a postdisciplinary body of knowledge. In turn, exploration and synthesis of postdisciplinary CPS knowledge assumes a combined use of collective interdisciplinary, multidisciplinary, and transdisciplinary research, while it does not reduce the role of unidisciplinary individual investigators. **Supradisciplinary research** emerged as a new doctrine of combining these research approaches from epistemological, methodological, and procedural perspectives. However, no methodology can be found in the literature that could facilitate the practical execution of supradisciplinary research programs and projects.

This issue has been addressed in the background research which has had a combined explorative and constructive nature. This chapter discusses the concepts of problematics, pluridisciplinary and postdisciplinary research approaches, their methodological and epistemological features, and operationalization of supradisciplinary research in the specific context of CPSs. It is based on the following published papers: H19, H20 and H21 (see Appendix A.1.1). In addition to the investigation of the fundamentals of pluridisciplinary

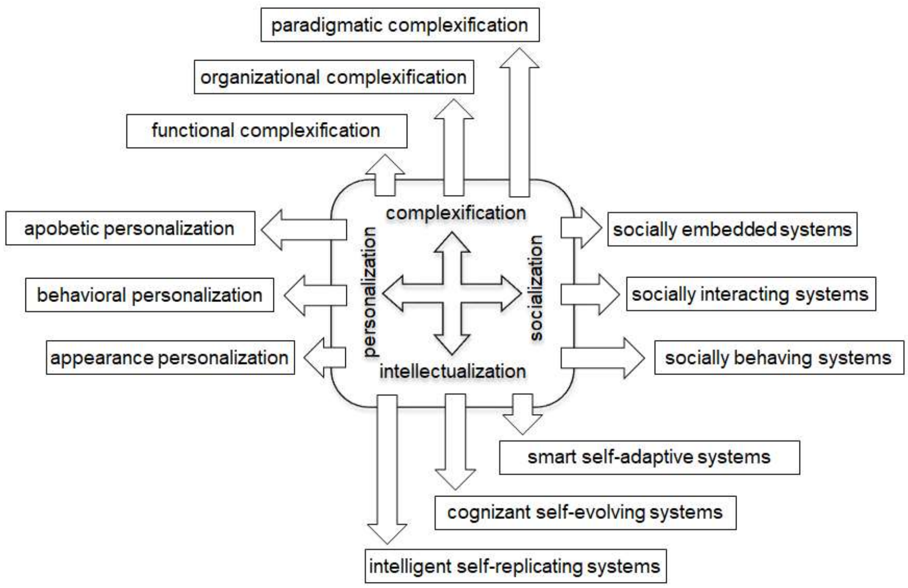

Figure 8.2: Recent trends in the evolution of CPSs

research approaches, the principles of **team science**, and other issues of conducting supradisciplinary research, a conceptual framework has also been proposed, which can be used as a blueprint for operationalization of such undertakings. The framework rests on six generic pillars: (i) problematics, (ii) infrastructure, (iii) methodics, (iv) stakeholders, (v) operations, and (vi) knowledge. It specifies the major concerns that have to be taken into consideration in a systematic manner at developing executional scenarios for supradisciplinary research. The framework arranges the concerns in a procedural logic - as they should be considered by the research managers and CPS developers.

8.3. Problematics of early stage collective research for cyber-physical systems

The above-described **multidimensional diversification** of CPSs raises many new challenges to cope with. In addition to the already known challenges (such as aggregative complexity, technological heterogeneity, functional dependence, stakeholder involvement, operational resilience, and safety and security), a partly research methodological and partly engineering epistemological challenge is also raised. There are two sources of this latter challenge. The first one is in the need to explore and synthesize proper cross-disciplinary (holistic) knowledge for conceptualization, design, and engineering of novel CPSs (Simon and Schiemer, 2015). The second one is in the traditional reductionist culture of system development that separates the tasks of system development and allocates them to distinct departments (Horváth, 2015). These two phenomena eventually boil down to the need for collective research and feeding the development process with synthesized disciplinary knowledge. In turn, this gives floor to the abovementioned research methodological and engineering epistemological challenge (Horváth, 2004). The two main questions are: 'How to conduct postdisciplinary research in the early stages of CPSs development?' and 'How to synthesize the knowledge during the research process towards the most reliable and useful shared intellect?

The main conjecture (and working hypothesis) underpinning the research has been that a pluridisciplinary or postdisciplinary research approach may fulfil the requirement. These inquiry approaches require two or more disciplines to combine their knowledge, methods, and expertise to jointly explore, confirm, and deliver research outcomes (e.g. theories, laws, facts) appropriate for a common subject area. The remaining part of the chapter argues that, in line with the multidimensional diversification of CPSs, a supradisciplinary research approach is indeed needed to explore and scrutinize knowledge for their development in the conceptualization and design stages. Though the concept of supradisciplinary research is known and addressed in the related literature, apart from the general methodological dispositions, no specific conceptual frameworks, framing methodologies, or process scenarios have been presented, in particularly not in the context of research activities and knowledge demands in the early developmental stages of cisp*CPSs (Balsiger, 2004). The author interprets it as a 'problematics' in itself, which involves intertwined scientific, technological, procedural, social, human, and business aspects.

The current literature is scarce on publications that would explain this 'problematics' and would offer a receipt for dealing with it. As posited in Horváth (2022), the science of CPSs is still in the stage of formation and the methodologies of doing pluridisciplinary research in this field have not yet reached further than their embryonic stage. Concerning the development of comprehensive multidisciplinary or transdisciplinary theories, only a few efforts have been reported in the contemporary literature (Horváth and Pourtalebi, 2015). Therefore, based on a synthesis of the outcomes of his previous research work, a multi-focal literature survey and expert interrogations, critical systems thinking and philosophical and methodological speculations, as well as on retroductive reasoning

(Ochara, 2022), the author proposes a generic framework for conducting supradisciplinary research in multidimensionally diversified CPSs. The importance of this work and the proposed framework lays in the foreseen even more intense diversification of the next generation of CPSs. The next section analyses the methodological convergence of the various research approaches. Afterward, the main tenets of team science are discussed, and the different forms of research co-working are presented. The last part of the chapter presents the six domains of concern and the blueprint of the proposed framework.

8.4. Convergence of individual and collective research approaches

The 21st century science is concurrently driven by transdisciplinary convergence, structural reorganization, and social transformation (NASEM (2019). The fact of the matter is that both convergence and divergence are perpetually present and interoperate in science, knowledge, and technology. Consequently, new competing research philosophies and strategies are emerging that have not been consolidated yet. The last decades have seen the transition from the so-called **Mode 1 science to Mode 2 science** (Nowotny et al., 2003). The former is the old paradigm of scientific inquiry and (i) is characterized by the hegemony of theoretical and experimental discoveries, (ii) establishes an internally-driven taxonomy of branches and disciplines of science, and (iii) acknowledges the autonomy of individual scientists and their host institutions. It is characterized by an analytical thinking approach that has its roots in reductionism (Horváth, 2017). The latter (i) is a new paradigm of socially distributed knowledge production, (ii) has a pluridisciplinary, collaborative, and application-oriented nature, (iii) is the subject of multiple accountabilities, and (iv) is typically considered in technological, social, political, and economic contexts. Analytic and prognostic systems thinking play an important role in practicing Mode 2 science. Systems thinking (i) explains the manifestation and behavior of systems as a whole, (ii) is dominated by abstraction and synthesis, and (iii) studies emerging and relational properties. It is supposed to be extended to the human behavioral domains (cognition, communication, leadership, etc.) (Horváth, 2023). As a consequence, systems thinking increases all forms of complexity and heterogeneity, but facilitates addressing sustainability of the environment, economy, and society. The overview of the latest theoretical and methodological developments in scientific research also needs (critical) systems thinking (Rousseau, 2017).

It is known that the terms interdisciplinary, multidisciplinary, and transdisciplinary are ambiguously defined and interchangeably used in the literature. Therefore, it seems to be useful to elaborate on the interpretation and use of these terms in this part of the Chapter. With this in mind, a comprehensive landscape of generic research approaches is shown in Figure 3. As an overall trend, the move from an individual focused unidisciplinary research approach, through pluridisciplinary ones, toward postdisciplinary approaches has been identified (Parker, 2008). In practice, it means that pluridisciplinary programs are also conducted in addition to monodisciplinary research programs. A research approach is "multiple disciplinary" if more than one discipline is involved, but the nature of their involvement is unknown or unspecified.

As a common term, pluridisciplinary refers to research that may involve interdisciplinary, multidisciplinary, and transdisciplinary research programs and approaches (Scholz and Steiner, 2015a and 2015b). While monodisciplinary research inquiries are conducted from the perspective of a single discipline, pluridisciplinary programs make attempts to investigate phenomena and problematics from multiple perspectives in an integrated manner. Supradisciplinarity is the descriptor of the doctrine of hybridization between knowledge domains. It means conducting monodisciplinary, interdisciplinary, multidisciplinary, and transdisciplinary research programs or activities simultaneously and purposefully in concert. While monodisciplinary research can be best characterized by the

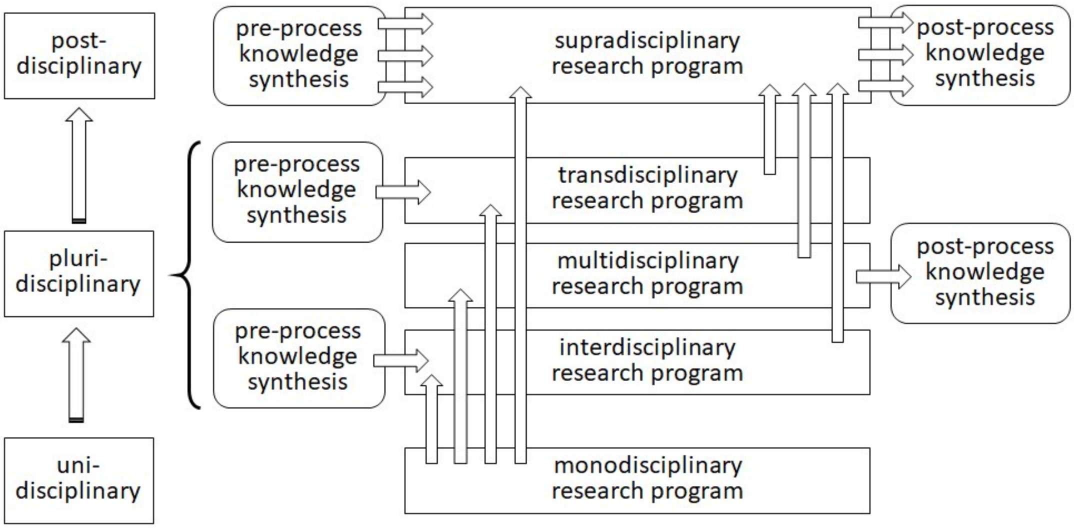

Figure 8.3: Overview of the generic research approaches

word ‘distributed’, a single-word descriptor for interdisciplinary research is ‘interactive’, for multidisciplinary it is ‘additive’, and for transdisciplinary it is ‘holistic’. As a recently conceptualized realization of a postdisciplinary philosophy, supradisciplinary research can be depicted as a ‘combinatorial’ approach.

The unidisciplinary approach practiced by individual investigators is a historical development. It counts on the insights and talents of individual researchers, working independently or in groups. Typically, doctoral (promotion) research regulations and frameworks still strictly follow the ideal of unidisciplinary, single-investigator research approaches. Notwithstanding, due to institutionalization of scientific research, it has been scaled up to large monodisciplinary projects based on team formation and collaboration. Though it is often deemed as exhausted from a praxiological perspective, it remains an indispensable kernel of doing scientific research. Neither modern pluridisciplinary nor postdisciplinary (or metadisciplinary) research philosophies are against this.

Interdisciplinarity assumes creating links between disciplines and a coordinated inquiry approach, including the establishment of a shared knowledge and method platform for launching projects (front-end integration) (Buanes and Jentoft, 2009). Interdisciplinarity is a subject of philosophical argumentation (Schmidt, 2008). Interdisciplinary research addresses phenomena that are not directly and completely covered by the concerned disciplines (Klein, 2010). It assumes close interaction in co-working and assumes the agreement of the investigators from different disciplines on the objectives and the different analysis and synthesis methods (CohenMiller and Pate, 2019). An interdisciplinary research approach involves the interaction and coordination between more than one discipline, aiming at (i) development of knowledge in each of the concerned disciplines, (ii) transferring knowledge from one discipline to another, and (iii) transforming knowledge of one discipline under the influence of another discipline (Fiore, 2008). These assume collaboration and the emergence of a new thought style (Darbellay, 2015).

Multidisciplinarity draws on knowledge from different disciplines, but stays within their boundaries (Alvargonzález, 2011). Multidisciplinary research projects are carried out independently by unidisciplinary researchers, but they are informed about the work of the other disciplines throughout the process. New knowledge is learned through the individual interest windows of the included disciplines, and evaluated and combined in the conclusive stage of research projects. In other words, the novel knowledge is synthesized and consolidated at the end of the conducted projects (back-end integration). No specific execution methodologies have been publicized for this purpose. Multidisciplinary research collaboration involves: (i) collective determination of the goals, (ii) working out a strategy/approach for achieving the goals, (iii) sharing physical, intellectual and intangible resources, and (iv) building common grounds and consensus. Multidisciplinary research teams are supposed to produce (i) a coherent picture of the subject matter of the scientific study, (ii) a description/explanation of (parts of) the problematics or phenomenon, and (iii) a set of ranked theories to underpin potential theories or solutions.

Transdisciplinarity integrates natural, social, and technical sciences in a social context and transcends their traditional boundaries (Bergmann et al., 2012). Transdisciplinarity also signifies lively interactions between stakeholders and crossing the boundaries between scientific and non-scientific communities (representatives of industry, government, and/or civil organizations) with the goal of reaching out to the entire society (Ashby et al., 2018). Practicing transdisciplinarity is usually a challenging task because it needs an epistemological and organizational framework (Giri, 2002). Transdisciplinary research is a blending of interdisciplinary perspectives to produce a hybrid perspective of two or more disciplines (Mobjörk, 2010). Researchers from different backgrounds have to find each other, get acquainted, and derive a common motivation (Brown, 2015). They should form linked research teams and research communities. A successful research conduct assumes an explicit specification of the goals as well as a widely-based knowledge and process synthesis (front-end integration) (Burgin and Hofkirchner, 2017).

That is, before the start of their collaborative work, the investigators have to synthesize a common platform of shared knowledge (Nicolescu, 2014). This knowledge is subject to new types of quality control and extra-scientific social criteria, including public review (Jahn and Keil, 2015). The researchers, who are co-working in teams and communities, must learn to understand and appreciate each other's perspectives, forerunning work, and new results. They should work out procedural and administrative scenarios for long term cooperation or coadunation. Without these, they cannot reap the extra benefits of collaborating across disciplinary boundaries. (Cilliers and Nicolescu, 2012). Working in a transformative manner has also been identified as a paradigmatic characteristic of transdisciplinary research programs and projects (Lawrence et al., 2022).

Supradisciplinary research is deemed a conceptually and empirically grounded constitutional element of Mode 2 knowledge production, or simply as if it was the same thing as Mode 2 science. It blends knowledge of more than one discipline and interlaces their research approaches. It assumes that the involvement of disciplinary investigators is designed, planned, and specified before launching any research program or project. Supradisciplinary research has ontological, epistemological, methodological, and praxiological conditions. The ontological condition is that its paradigm is accepted as trustful and realistic. The epistemological condition is that a preliminary knowledge synthesis (Hoffmann et al., 2017), as well as an unbiased synthesis of the novel findings, is possible (Defila and Di Giulio, 2015). The methodological condition is that a conceptual framework/platform can be created together with a pool of complementary research methods. The praxiological condition is about overstepping the (research) cultural boundaries by a holistic relational process, in which knowledge is produced through transactions of stakeholders and a supporting communal project management (Tebes et al., 2014). Consequently, the knowledge created by supradisciplinary research is understood to be: (i) reflexive with regard to social accountability, (ii) traceable back to its starting point in societal needs, and (iii) servicing its social 'stakeholders' and applied research. In other words, it offers a genre of knowledge that manifests in the contexts of applications and, therefore, cannot be classified according to a distinction between branches.

8.5. From phenomena to problematics and from problematics to phenomena

The term 'problematics' is eventually an abstraction. It is used to refer to the existence of multiple, extensive, holistically evolving and interacting practical research challenges that are (i) uncertain or not settled, (ii) complicated to handle and solve, (iii) difficult to decide upon, and (iv) their future state is open to debate (Osborne, 2015). The essence of problematics is a set of inherently intertwined wicked scientific problems of the same nature, which appear, for instance, in the demand for sustainable living or in the uncontrolled proliferation of artificial intelligence (Pohl, 2005). Other examples of these multi-factorial challenges are such as climate change, energy provisioning, circular production, extreme social stratification, ecological sustainability, informational smog, well-being, pandemics and chronic diseases, profitable recycling, socio-technological problematics cannot be reduced to component problems due to their innate holism.

From the perspective of making systematic inquiries, all problematics are associated with and caused by the interaction of a number of underlying natural phenomena and artificial (synthetic) phenomena. As a simple example of a natural phenomenon, we may consider the CO_2 absorbing capability of plants and the $C0_2$ production of vehicles powered by internal combustion engines as an artificially created phenomenon. The interaction of these phenomena creates an 'everyday problematics' that manifests itself as air pollution. Like capturing natural and artificial phenomena, a comprehensive description, explanation, prediction, or regulation of problematics also requires evidence-based scientific theories. However, establishing these theories is a complicated matter because of the reason that the involved phenomena (or parts thereof) are usually studied within different disciplines. Technologically, economically, or societally created problematics need collective scientific inquires and novel research approaches (such as computational modeling-based holistic simulations and multi-level massive data processing).

8.6. Team science to assist the formation of Mode 2 science

According to a widely accepted definition, team science (TS) is a collaborative effort to address a scientific challenge that leverages the strengths and expertise of professionals trained in different fields (NRC, 2015). TS is one of the set of strategies and efforts that advance convergence (Ledford, 2015). It utilizes the core principles and best practices of community psychology to enable the current transformation in science and to develop research competencies for groups and communities. One of the main objectives is to develop general principles and trustworthy effective practices for both co-located and dislocated (online) co-working (Mâsse et al., 2008). TS is not against the traditional single-investigator driven research approaches, but wants to learn from their limitations and augment them towards group- and community-oriented research approaches. It studies how coordinated teams of diverse skills and knowledge can tackle complex scientific and societal problematics and emergent issues (Boardman and Ponomariov, 2014). It promotes both intra-personal competencies (attitudes, knowledge, skill, experiences) and inter-personal competencies (socialization, communication, empathy, trust). TS also studies elements of team and community processes (common vision, communal mission, tactical goals, shared understanding, responsible roles) and institutional infrastructure and policies (hiring, promotion, experiences, desires, interdependence, organization, funding opportunities, data management, networking, road mapping) (Stokols et al., 2008). Furthermore, it studies the broader influences of co-working (e.g., history, cooperation framework, publishing forums, academic events, and industrial relations) (Vogel et al., 2013).

Undertaking research in a collaborative way is called upon by convergence. TS intends a generic theory to explain multi-scale collaboration in a level- and context-independent manner (Yu et al., 2019). The primary conjecture is that a research team- or community-based approach is the proper organizational form and one key strategy for tackling complex problems across boundaries. However, it is challenged by the difficulties of establishing partnerships across multiple local and international institutions. It remains a task for top managements of these institutions to recognize, institutionalize, and operationalize team science. They have to think of both horizontal integration (bringing together disciplines that share common features, methodological approaches, and overlapping background knowledge) and vertical integration (linking disciplines across multiple types and levels of analysis and synthesis). Insights from the social and behavioral sciences help form and sustain effective research teams and communities. In addition, bridging across interrelated research interest areas is also supported by international personal networks and individual investigators, who have broad expertise in more than one area and embody the idea of convergence.

Convergence transcends disciplinary boundaries, even extending beyond what is traditionally regarded as science. Reportedly, (i) raising public/professional awareness of convergence, (ii) building common grounds and consensus, and (iii) establishing scientific cultures that support convergence are catalysts of new scientific knowledge and applications. Formation of integration (from incidental partnership to strategic alliance) is a nondeterministic, interest-driven process. In current practice, integration of collective work is usually emergent and volatile, driven by project calls and the interest of funders. Theories of team science should be more articulated with regard to the various practical manifestations of collective work, which can be: (i) **cooperation** (involving information sharing and supporting organizational research outcomes), (ii) **coordination** (harmonizing research activities and support of mutual benefits), (iii) **collaboration** (giving up some degree of research independence in an effort to realize a shared goal), and (iv) **coadunation** (achieving the state or condition of being united by gradual synergy forming and growth in research) (Figure 8.4). Theories must also explain the time-dependence of the drivers (e.g., temporal changes in complex societal needs, advancement of technologies, novel business models, diversification of knowledge, and time-influenced organizational principles.) and the obstacles of co-working (e.g., culture of coping alone, attitudinal disinterest, IP protectionism, insufficient competencies, fear of transparence).

Scientific research has been heavily institutionalized over the last century and typically conducted in hierarchical organizations. Its structural reorganization proceeds according to the concept of organizational heterarchy that establishes interdependence, even independence, relationships of the stakeholders of research. In addition to extensive academic research collaboration, addressing complex societal challenges also needs collaboration with various public stakeholders (Melo and Caves, 2020). The paradigmatic model has been multi-institutional project-based research collaboration in cooperating teams. In such projects, collective competence and wisdom has been deemed more essential than individual ingenuity and diligence. Typically experienced in research cooperation over geographic, economic and cultural boundaries, the

Figure 8.4: Forms of co-working in pluridisciplinary research

term 'cultural diversity' has been used widely to refer to the differences of humans/societies in a specific region. It is reflected in the mental models and behavioral styles, and results in different value systems. Cultural diversity in research is a difficult matter to deal with, because reasoning, decision making, behavioral and interaction models are all involved. The literature also informs us about the fact that many multi-national collaborative research projects have suffered from 'research cultural clashes' in the lack of a common ground, building awareness, showing openness, and exercising patience.

8.7. Related organizational, management, and social issues

Pluridisciplinary research approaches consider some observable problematics (or phenomena) from different disciplinary viewpoints and intend to neatly merge and integrate relevant parts (concepts, models, methods, findings) of different scientific disciplines in systematic inquiries (Lawrence, 2015). These approaches require understanding the previous research and perspectives of the others with whom they intend to work and go together with a confrontation of the scientific concepts, models, methods, and findings of the concerned disciplines. They are epistemologically and socially constructivist approaches that highlight the emergent, recursive, and communal nature of knowledge production. The various approaches have different positions with regard to integration, and attempts have been made to develop formal composition methodologies and methods for fusing harmonizing theories (Wiek, 2007). They advocate that much knowledge emerges through academic interaction during the research process and through social action itself and recognize informal social and cultural integration mechanisms as essential. The need for synthesizing and sharing knowledge is present and important during the early negotiation phases of research projects in order to build up the required redundancy, as well as at the conclusion phases of research projects in order to consolidate the new knowledge (Wiesmann et al., 2008).

The establishment process of supradisciplinary research programs and projects offers opportunities for addressing social issues and enhancing the social skills of the involved researchers representing multiple disciplinary domains. The specific goals of **social management** are visualized in Figure 8.5. Stimulated by the work of (Tebes and Thai, 2018), this model considers six stages of academic and public socialization of research. The goal in the first stage is to learn and interlink the subject knowledge, methodological know-how, and working culture of the researchers representing the involved competence domains. The goal of the second stage is sharing mental models and conceptual frames, whereas the goal of the third stage is exchanging interdisciplinary synthetic skills and experiences. These two processes contribute to the development of **collaborative social skills**. The specific goals of the fourth stage are to establish a hybrid virtual knowledge base and a virtual method warehouse, with the intension of helping the involved individual researchers and research teams to familiarize themselves with the bodies of knowledge and the arsenal of methods, tools, and instruments used by others. Together with activities in the two preceding stages, this creates a so-called professional reference space for all researchers and teams. In the fifth stage, a **joint problematics space** is created and maintained. This makes it possible for the involved researchers to address the same problematics, while they look at and interpret it from their own perspectives. As an outcome, epistemic translations take place which offer deeper insights and blur the boundaries if they still exist in the **professional reference space**. The last stage is the actual organization and management of supradisciplinary research programs and projects, which is facilitated by the growing social awareness and managerial competencies.

Pluridisciplinary research approaches direct attention to dimensions such as research management, partnership, sharing, productivity, and exploitation (Hadorn et al., 2008).

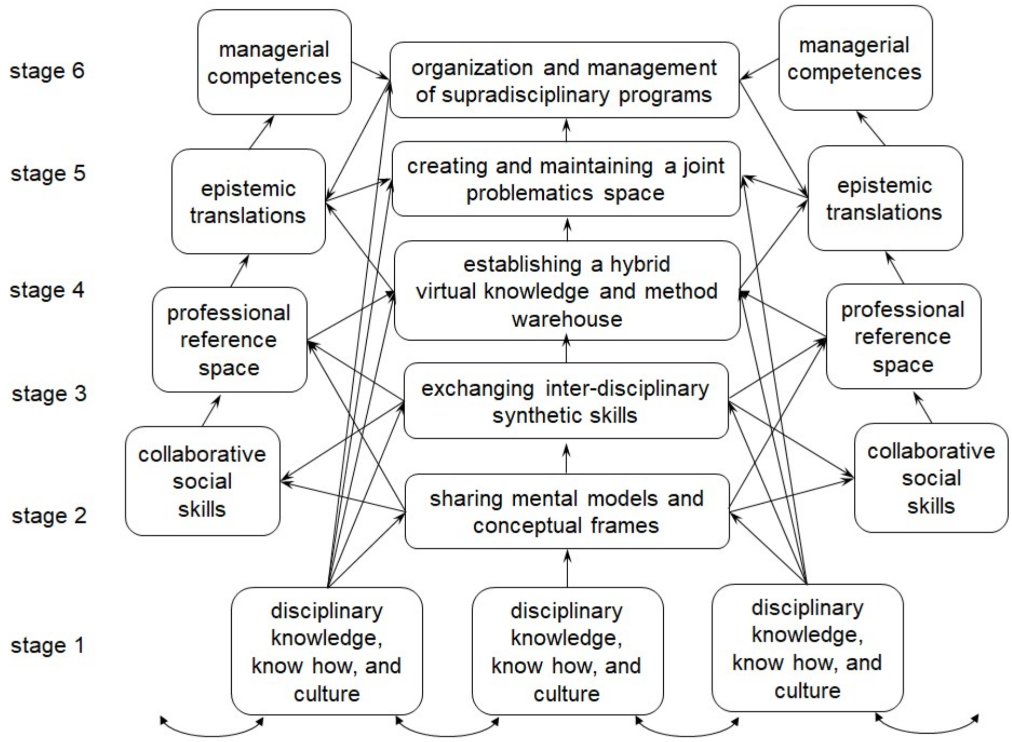

Figure 8.5: Creating joint intellectual spaces

Actually, they place the layers of organization, management, and utilization above the layers of competence development, execution of inquiry, and dissemination of the results. Successful program/project management also assumes exposing academic leadership. In general, research leaders should (i) address the barriers of effective professional and social convergence, and strong partnerships within and across institutions, (ii) develop policies, guidelines, and practices to support and evaluate convergent research, (iii) utilize the expertise of economic, social, and behavioral sciences to realize best practices, (iv) master program management and strategic planning when forming a research initiative, (v) be aware of the most effective recruitment practices, research support policies, risk analysis and recovery models, cost and revenue allocation models, including catalytic seed funding's, and (vi) apply comparative evaluation policies, tenure and promotion advancement, and unique evaluation criteria for rewarding both professional and social achievements.

8.8. Domain of concerns and a blueprint of the proposed framework

Organization of problematics-driven design research for the development of next generation CPSs is a new challenge. Due to its complexity, it can be addressed only by a supradisciplinary research approach that enables collective knowledge exploration and integration processes. As discussed in the preceding sections, such an approach is influenced by a large number of factors. Of paramount importance are: (i) the organization theory of holistic co-processes, (ii) the principles and recommendations of team science, (iii) the praxiological issues of 21st century (Mode 2) science, (iv) the societal epistemology of emerging problematics, and the (v) psychologic theory of creative communities (Figure 8.6). Operationalization of a supradisciplinary research approach needs a conceptual framework, which is supposed to specify both the ontological pillars and the methodological-procedural concerns. Based on what has been discussed in the preceding sections of the dissertaion, six ontological pillars have been identified, namely: (i) the investigated complex problematics (or phenomena), (ii) the integrated and shared research infrastructure, (iii) the applied research methodics, (iv) the involved academic and public stakeholders, (v) the establishment and execution inquiry operations, and (vi) the input and output knowledge. Shown in Figure 8.7, these ontological pillars of supradisciplinary research are actually strongly interlinked, even intertwined.

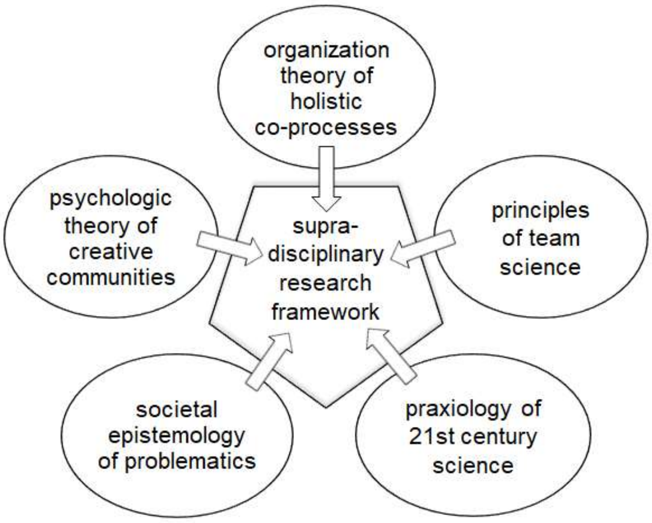

Figure 8.6: Main factors influencing a framework of supradisciplinary research

In simple words, the main requirement concerning the conceptual framework is to explain what the associated concerns are, and in which order they should be taken into consideration at organizing supradisciplinary research (McComb and Jablokow, 2022). As the upright pillars supporting a building are made up of bricks, the ontological pillars of the framework are built from conceptual building blocks. In fact, these building blocks are concerns of realization. This means that the structure of the conceptual framework has been defined by the ontological pillars, whereas its specific content has been compiled from the related concerns of realization.

The blueprint of the conceptual framework is shown in Figure 8.8. Based on their intertwined nature, no ordering logic can be imposed on the ontological pillars. On the other hand, the order of the blocks of the pillars shows a procedural order. For instance, selection of the problematics for study commences with exploring the space of relevant complex problematics (or phenomena) and finding the most appropriate one (first choice) for common research interest, and concludes with planning and starting the exploitation of the novel intellectual assets. The establishment of a shared research infrastructure starts with the overview of the existing research facilities and the planning of the additionally needed facilities towards a functional integration of laboratories, and finishes with designing and provisioning security and safety services. The elaboration of the research methodics (designing the inquiry processes, defining sets of research methods and instruments, and stating the applicability and performance criteria, without a common underpinning theory) starts with a research task analysis and concludes with coherence and correspondence analysis of the outcome theories.

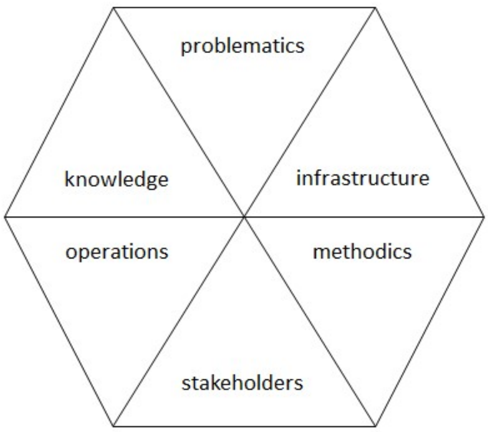

Figure 8.7: Six pillars of implementation of supradisciplinary research

The **stakeholder integration** starts with partnering strategy development, involving both academic and public stakeholders, and concludes with the enhancement of synthetic professional and social skills. The establishment and execution of inquiry operations starts with the elaboration of the principles of efficient operative leadership and extends to the management of progress reporting and reviews. The engineering and management of input and output knowledge commences with studying the past activities and background knowledge of the partners, and concludes with the

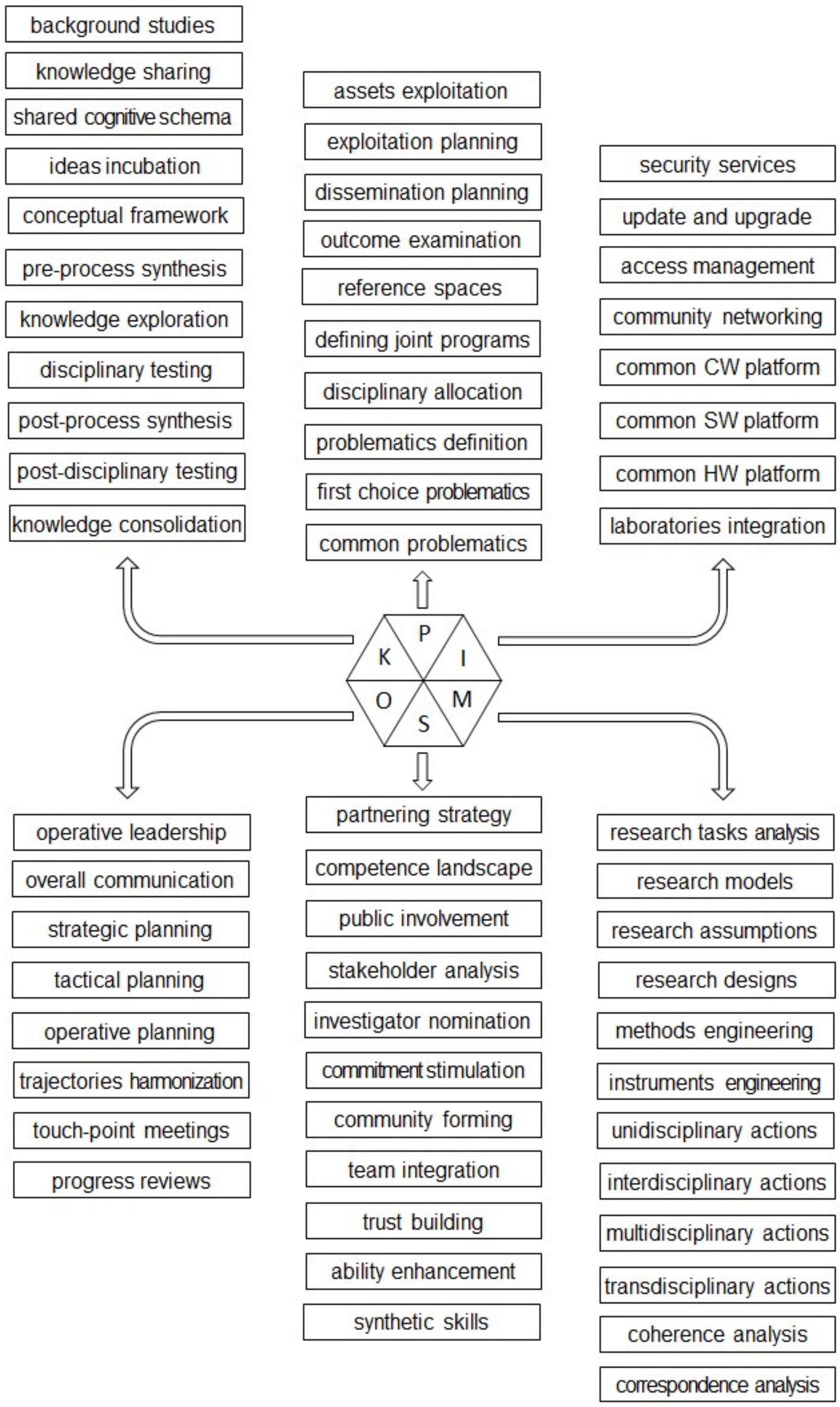

Figure 8.8: The blueprint of the proposed supradisciplinary research framework

internal and external consolidation of the new knowledge. In practice, working on the six ordered groups of concerns happens concurrently and interdependently.

The proposed framework advises on how to organize and execute supradisciplinary research, but it does actually not describe what to study. Therefore, it should be seen as a kind of meta-research design that has to be combined with specific research models that are foreseen results of addressing the dedicated concerns in the thread of methodics. An associated research model or multiple integrated research models can convey information about complex problematics or phenomena that have significance and do indeed need supradisciplinary research approach, including team- and community-based inquiry efforts.

8.9. Reflections and possible follow up research

A novel narrative is emerging for 21st century science that believes in interpersonal transactions while working in teams and communities, as well as in active engagement of public stakeholders by researchers in research programs or projects addressing socially-based problematics (Kläy et al., 2015). The objectives of these postdisciplinary research approaches are to: (i) holistically investigate and resolve complex problematics of the real world, (ii) provide different perspectives on and approaches to such complex problematics, (iii) offer holistic theories to answer research questions posed by multiple disciplines, (iv) develop consensus about definitions, principles, and guidelines to deal with non-reducible complicated systems, and (v) provide novel and comprehensive services for knowledge exploration and synthesis. In this chapter, I have discussed the concepts of pluridisciplinary and postdisciplinary research approaches, their methodological and epistemological features, and operationalization of supradisciplinary research in the specific context of CPSs.

- Supradisciplinary research opens up a multidimensional space of inquiry that is characterized by (i) concurrent dependence on multiple (physical, biological, human, social, computational, technological, etc.) domains of inquiry and investigations, (ii) various progression levels (discovery, description, explanation, prediction, and regulation) with regard to the studied phenomenon (problems), and (iii) need or synergy in terms of hardware, software, cyberware, brainware, etc. related knowledge.
- Doing supradisciplinary research is not trivial due to the growing overall complexity, functional and technological heterogeneity, and knowledge-driven nature of engineered systems. It has been learnt that acceptance of transdisciplinary research is negatively influenced by (i) the growing versatility of professional knowledge and complexity of academic cooperation, (ii) the historical compartmentalization of the scientific landscape, (iii) the sectoral division of responsibilities in contemporary society, and (iv) the increasingly diverse nature of the societal contexts.
- Based on informed assumptions, a six-pillar conceptual framework has been proposed. It clarifies the concerns associated with the establishment and execution of community-based supradisciplinary research programs/projects. Though the framework has been developed with a view to the specific application field, it is general enough to be transferable to other similar fields with or without adaptation.
- The framework rests on six generic pillars: (i) problematics, (ii) infrastructure, (iii) methodics, (iv) stakeholders, (v) operations, and (vi) knowledge. It specifies the major concerns that have to be taken into consideration in a systematic manner at developing executional scenarios for supradisciplinary research. The framework arranges the concerns in a procedural logic - as they should be considered by the research managers and CPS developers.
- Though its importance is recognized, in its current form, the framework does not cover the specific societal and personal issues of a successful organization of the inquiry at individual researchers, research teams, and research community levels. Notwithstanding, the framework can facilitate (i) management of research program and project organization tasks, (ii) joint formation of shared research infrastructure, (iii) setting up concrete research programs, projects, and processes, (iv) academic partnering and public stakeholder involvement, (v) process flow management and capacity/competence allocation, and (iv) knowledge synthesis, assessment, and consolidation in a holistic manner.

- Follow-up community-based research may focus on (i) the practical application and testing of the framework in concrete cases – a task that an individual researcher cannot address, (ii) the general principles and concrete methods of organizing effective knowledge sharing and long term postdisciplinary work of multiple research teams or a research community.
- There is a paradoxical situation with the research reported in this chapter. The source of the paradox is the juxtaposition between the form of the inquiry used to produce knowledge for the cognitive framework and the form of the work that would have been needed for a full-scale operationalization of the framework in practice and to validate it through its implications over multiple application cases. The inquiry could be done by one individual investigator who has sufficient knowledge of multiple disciplines (domains of interest) such as research theory and methodology, engineering and technologies of CPSs, and systems science and thinking (Dawson, 2013). However, projection of the contents of the conceptual framework to real life problematics and completion of the necessary supradisciplinary research would need multiple collaborating research teams or even a research community. Therefore, execution of the activities proposed by the conceptual framework cannot be done and the contents and implications of the proposed framework have not been rigorously scrutinized.

Chapter 9

9. Epilogue

9.1. Scientific propositions of the dissertation

The reported research has been done to enrich knowledge and improve the approaches of **cognitive design and engineering** of cyber-physical systems. This emerging disciplinary domain is interested in the problem-solving intellect of this family of systems, and intends to explore knowledge, methods, and tools to support their intellectualization processes and synthetic intellect-based application-specific operations (Srinivasan, 2013). Though it has some overlap with artificial intelligence research and reuses parts of the knowledge and means developed by intelligence researchers, the above objectives demarcate it from that. Cognitive engineering addresses emerging, complicated, prognostic, but influential general subject matters, rather than issues related to specific system implementations and applications. It focuses on facilitation of the development of next-generation CPSs, rather than on systems reproducing parts and forms of human intelligence. It is unconcerned with the dilemma if human intelligence can, may, must, or will sooner or later be replicated (Bostrom, 2014). But it is concerned with how people can benefit from the services of complexificated, intellectualized, socialized, and personalized systems. Consequently, the presumed aim of the reported research was, on the one hand, to contribute to the disciplinary domain of cognitive engineering and, on the other hand, to explore probabilities and possibilities for generating benefits. Based on the completed work and the obtained results, the following scientifically significant propositions could be stated:

Proposition 1:

The use of paradigmatic system features and paradigmatic feature profiles for a generic characterization of existing CPSs or prescriptive specification of new ones is preferable to textual, visual or symbolic formulation of augmentative, descriptive, normative, predictive or domain-specific definitions. Δ

Proposition 2:

The implementation of smart CPSs requires a computational model that brings together: (i) sensing and monitoring based on multiple sources, (ii) awareness of the state of the system, the environment, and the problems, (iii) ampliative reasoning based on the system's knowledge, (iv) strategic learning from the processes and results, (v) evaluation of the actual and planning the optimal operation, (vi) runtime planning and validation of the system's adaptation, and (vii) actuating the modified generative, transformative, and informative core-functions, and that complements the latter with parafunctions such as (i) sociality, (ii) personality, (iii) ingenuity, (iv) dexterity, (v) convincingness, and (vi) dependability, which are indicators of the system's smart behavior. Δ

Proposition 3:

The successive generations of CPSs can be differentiated based on their (interrelated) level of self-intelligence and level of self-organization and, according to the introduced paradigm advancement model, they can be identified as zero, first, second, third or fourth generation. Δ

Proposition 4:

The system-level synthetic problem-solving knowledge acquired and/or produced by intellectualized CPSs represents a unique genre of knowledge that must be distinguished from the known alpha, beta, gamma, and delta (ABGD) genres of knowledge. This epsilon-knowledge is becoming a fully-fledged complement of the ABGD genres of knowledge. Δ

Proposition 5:

Being not addressed directly by gnoseology or epistemology, scientific investigation of system-level synthetic problem-solving knowledge requires a new theoretical foundation, methodological approach, and investigational structure. The concept of sympérasmology has been outlined to provide these means and to address four specific study areas, namely (i) the fundamental notions, (ii) the basic principles, (iii) the capabilities, and (iv) the implications, each of which is broken down into multiple interrelated aspects of investigation. Δ

Proposition 6:

The elaborated prognostic systems thinking approach, which supplements the conceptual pillars of traditional analytical systems thinking with prognostic conceptual pillars in a framework, interpreted as a semantic network, and which supports forward-looking examination of systems with a structured and expandable library of analysis aspects, is indispensable from the point of view of the prognostic examination of internal and external changes, relationships, and effects of dynamically evolving systems. Δ

Proposition 7:

System-level synthetic problem-solving knowledge can be shared among and exploited by similar CPSs if (i) the knowledge transfer methods which are associated with knowledge-bases/repositories, implemented by software agents, used in transfer/distributed learning, and belonging to knowledge graphs/collective intelligence are integrated and necessarily supplemented, (ii) the system-level synthetic problem-solving knowledge and the related reasoning mechanisms are included in uniform transfer packages, and (iii) the transfer process is broken down into local neutralization, global transfer, and local naturalization sub-processes. Δ

Proposition 8:

In order to effectively explore, describe, and explain complex phenomena and problematics which are associated with conceptual design and implementation of next-generation CPSs, investigations going beyond the disciplinary boundaries are needed whose organizational, executional, and managerial framework is to be provided by the methodology of community work-based, societally-orientated, and epistemologically holistic supradisciplinary research. Δ

Proposition 9:

The ontological basis of supradisciplinary research conducted in the context of intellectualized, socialized, and personalized CPSs is formed by (i) the investigated phenomena or problematics, (ii) the associated underpinning knowledge, (iii) the shared research infrastructure, (iv) the complementary research activities, (v) the overall research methodics, and (vi) the stakeholders of research, and, by taking these into account, shared intellectual spaces can be created and specific collective implementation scenarios can be developed. Δ

9.2. Own reflections on the work and results

The overwhelming majority of the research work presented in the dissertation has been completed within the last six years. The results of some of the efforts have been released for public debate in international journals only in the last three years. The latest publications are still in 'print online' status. Whereas the earlier publications have been cited time proportionally, this can only be expected to happen for the most recent publications. Obviously, the external interest in them depends not only on their freshness, but also on the contributed novelty, utility, and implications of the knowledge they are delivering to the public, in addition to their academic and practical values. Interests may be expected for the reason that cognitive engineering of CPSs belongs to two of the five challenges of systemic digitalization that, according to (Ochara, 2023), includes (i) the circular economy, (ii) cyber-physical systems, (iii) sharing economy, (iv) digital transformation, and (v) smart systems. Whereas Industry 4.0 strategized the evolution of productive and servicing industries toward an extensive use of CPSs and to optimize their output thereby, the maturing idea of Society 5.0 concerns the necessary evolution of society toward sustainable socio-cyber-physical systems (Miller, 2022). In addition, intellectualized CPSs have been forecasted to play a growing role in education both as supporting means of self-managed autonomous learning and as a target of transdisciplinary systems education and professional training (Tang et al., 2020).

With regards to novelty, my subjective reflection is that most of the presented research work is related to phenomena or problematics whose investigation is still in an early stage, no matter if analytic explorative studies or constructive knowledge synthesis studies are concerned. In the case of several issues, it was not possible to find publications on the exact topic or on largely similar approaches at the beginning of my work. Since that, remarkable progress has been achieved in certain themes, while much less in others, though their importance is being emphasized. Many of the investigated topics (such as paradigmatic system features, cognitive engineering of systems, smart and intelligent operation of CPSs, application-specific problem solving knowledge and computational reasoning mechanisms, prognostic systems thinking, and reuse of SSK), are directly associated with (frontline) issues which do not have direct precedents. The reflections of manuscript reviewers and experts received so far indicate that the presented research work is considered forward-looking even in an international context, as it extends the boundaries of knowledge to important new areas and creates new innovation opportunities. Notwithstanding, further extensive studies of the topics discussed are equally important from the point of view of scientific cultivation, innovation stimulation, industrial development and enhancement of social welfare.

More specifically, the novel scientific contributions of the dissertation can be summarized as follows: (i) introducing and using paradigmatic system features profiles for identification and characterization of CPSs, (ii) constructing a self-intelligence and self-organization-based model of generational evolution of CPSs, (iii) distinguishing system-level synthetic problem-solving knowledge (SLS-PSK) from other genres of (human) knowledge, (iv) proposing sympérasmology for systematic investigation of SLS-PSK, (v) specification of a conceptual framework, constructs, and means for enabling prognostic systems thinking, (vi) developing a theoretical underpinning and computational methodology for aggregation and utilization of system-level intellect as an industrial asset, and (vii) framing supradisciplinary collective research for CPSs. Contrary to these, every research action has remained an open story in my eyes. This impression is with me not only because of the diversity and broadness of the phenomena and problematics studied, but also as a consequence of the many things learnt and recognized during the focused inquiry processes.

Regarding the implications of the knowledge shared in the dissertation, two possible areas of direct utilization can be identified. One of these is the inclusion of the systematized conceptual fundamentals, models and frameworks in the development of a consistent and coherent disciplinary educational curriculum. The other utilization possibility is the integration of the results into the design methodology of human- and system-centered complex, intellectualized, socialized and personalized CPSs. These are expected to lay the foundation for a broader innovation, and through this, will have an impact not only on the relevant sector of the industry, but also on people's everyday lives. I believe that the thesis and the dissertation contribute to both in a balanced way and stimulate further efforts. Important to note that the research results and the published papers related to the fundamentals, overview, and conceptualization of cyber-physical systems have proved to be very useful in regular education (more precisely, in the M.Sc. elective course titled Cyber-physical systems design, as well as in discipline-related Ph.D. courses) in the period from 2013 to 2020, but also in occasional on-line courses. With regard to the earlier published research results, beyond the academic reactions, the industrial impact and the extent of utilization of the results are not directly measurable yet.

Appendix 1

A.1. References cited in the dissertation

A.1.1. Own references in the order of being processed in the dissertation

H1 Horváth, I. (2012). Beyond advanced mechatronics: New design challenges of social-cyber-physical systems. In: *Proceedings of the 1st Austrian Center of Competence in Mechatronics Workshop on Mechatronic Design*, ACCM, Linz, Austria, 2012, pp. 1-20.

H2 Horváth, I. (2014). What the design theory of social-cyber-physical systems must describe, explain and predict? In: *An Anthology of Theories and Models of Design*, Springer, London, pp. 99-120.

H3 Horváth, I. (2018). Observation of intention, action and conduct by cyber-physical systems in home care context. In: *Proceedings of the 12th International Symposium on Tools and Methods of Competitive Engineering - TMCE 2018*, Delft University Publisher, Delft, pp. 75-92.

H4 Horváth, I., & Wang, J. (2015). Towards a comprehensive theory of multi-aspect interaction with cyber physical systems. In: *Proceedings of the International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, Vol. 57052, ASME, pp. 1-14.

H5 Horváth, I., & Gerritsen, B.H.M. (2012). Cyber-physical systems: Concepts, technologies and implementation principles. In: *Proceedings of the International Tools and Methods of Competitive Engineering Symposium - TMCE 2012*, Delft University Press, Delft, pp. 19-36.

H6 Horváth, I., & Pourtalebi, S. (2015). Fundamentals of a mereo-operandi theory to support transdisciplinary modeling and codesign of cyber-physical systems. In: *Proceedings of the International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, Vol. 57045, ASME, paper V01AT02A005, pp. 1-14.

H7 Horváth, I. (2021). Connectors of smart design and smart systems. *Artificial Intelligence in Engineering, Design, Architecture and Manufacturing*, Vol. 35, No. 2, pp. 132-150.

H8 Gerritsen, B.H., & Horváth, I. (2012). Current drivers and obstacles of synergy in cyber-physical systems design. In: *Proceedings of the International Design Engineering Technical Conferences & Computers and Information in Engineering Conference*, ASME, pp. 1277-1286.

H9 Horváth, I., Rusák, Z., & Li, Y. (2017). Order beyond chaos: Introducing the notion of generation to characterize the continuously evolving implementations of cyber-physical systems. In: *Proceedings of the International Design Engineering Technical Conferences & Computers and Information in Engineering Conference*, ASME, pp. 1-14.

H10 Horváth, I. (2022). Designing next-generation cyber-physical systems: Why is it an issue? *Journal of Integrated Design and Process Science*, Vol. 26. No. 3-4, pp. 1–33.

H11 Horváth, I. and Gerritsen, B.H.M. (2013). Outlining nine major design challenges of open, decentralized, adaptive cyber-physical systems. In: *Proceedings of the International Design Engineering Technical Conferences & Computers and Information in Engineering Conference*. ASME, Vol. 55867, pp. 1-14.

H12 Horváth, I. (2022). The epsilon-knowledge: An emerging complement of Machlup's types of disciplinary knowledge. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*. Vol. 36, paper e18, pp. 1-18.

H13 Horváth, I., & Duhovnik, J. (2005). Towards a better understanding of the methodological characteristics of engineering design research. In: *Proceedings of the International Design Engineering Technical Conferences and Computers and Information in Engineering Conference* (Vol. 47403), ASME, pp. 905-919.

H14 Horváth, I. (2020). Sympérasmology: A proposal for the theory of synthetic system knowledge. *Designs*, Vol. 4, paper 47, pp. 1-24.

H15 Horváth, I. (2020). On reasonable inquiry and analysis domains of sympérasmology. *Journal of Integrated Design and Process Science*, Vol. 24, No. 1, pp. 5-43.

H16 Horváth, I. (2022). A first inventory of investigational concerns for prognostic systems thinking based on an extended conceptual framework. *Journal of Integrated Design and Process Science*, Vol. 27, No. 1, pp. 1–25.

H17 Horváth, I. (2016). Theory building in experimental design research. In: *Experimental design research: Approaches, perspectives, applications*, Springer, Berlin Heidelberg, pp. 209-231.

H18 Horváth, I. (2022). Utilization of synthetic system intelligence as a new industrial asset. *Journal of Integrated Design and Process Science*, Vol. 26. No. 3, pp. 1-21.

H19 Horváth, I. (2015). An initial categorization of foundational research in complex technical systems. *Journal of Zhejiang University - Science A: (Applied Physics & Engineering)*, Vol. 16, No. 9, pp. 681-705.

H20 Tanik, M.M., Gatchel, S., Horváth, I., Wan, T., Kim, K. Y., Huang, J., ... & Zeng, Y. (2021). Footsteps towards a transdisciplinary design and process science. *Journal of Integrated Design and Process Science*, Vol. 25, No. 3-4, pp. 1-16.

H21 Horváth, I. (2023). Framing supradisciplinary research for intellectualized cyber-physical systems. *Journal of Computing and Information Science in Engineering*. Vol. 23. No. December 2023, pp. 060802-1 - 060802-10.

A.1.2. References related to Chapter 1.

Baier, M.S., Berger, S., Kreuzer, T., Oberländer, A.M., & Röglinger, M. (2023). What makes digital technology? A categorization based on purpose. *Communications of the Association for Information Systems*, 52(1), 4.

Baldassarre, G., & Granato, G. (2020). Goal-directed manipulation of internal representations is the core of general-domain intelligence. *Journal of Artificial General Intelligence*, 11(2), pp. 19-23.

Commuri, S., Albus, J.S., & Barbera, A. (2018). Intelligent systems. In: *Autonomous Mobile Robots*, CRC Press, Boca Raton, pp. 665-696.

Couch, J.R. (2023). Artificial Intelligence: Past, Present and Future. *Journal of the South Carolina Academy of Science*, 21(1), 2, pp. x-x.

Funtowicz, S.O., & Ravetz, J.R. (1993). Science for the post-normal age. *Futures*, 25(7), pp. 739-755.

Gill, H. (2006). NSF perspective and status on cyber-physical systems. Austin. Internet: http://varma.ece.cmu.edu/CPS/Presentations/gill.pdf.

Giri, A. (2002). The calling of a creative transdisciplinarity. *Futures*. 34. pp. 103-115.

Gosling, W. (2020). Drivers of technology. In: *Culture's Engine*, Palgrave Macmillan, Singapore, pp. 61-65.

Gottfredson, L.S. (1997). Mainstream science on intelligence: An editorial with 52 signatories, history, and bibliography. *Intelligence*, 24, pp. 13–23.

Horváth, I., & Pourtalebi, S. (2015). Fundamentals of a mereo-operandi theory to support transdisciplinary modeling and codesign of cyber-physical systems. In: *Proceedings of the International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, Vol. 57045, ASME, p. V01AT02A005, pp. 1-14.

Lungarella, M., Iida, F., Bongard, J.C., & Pfeifer, R. (2007). AI in the 21st century - With historical reflections, Springer, Berlin Heidelberg, pp. 1-8.

Miki, T., & Yamakawa, T. (2008). Beyond the brain-inspired information technology. In: *Proceedings of SCIS & ISIS 2008*, Japan Society for Fuzzy Theory and Intelligent Informatics, pp. 1774-1777.

Monett, D., & Lewis, C.W. (2020a). Definitional foundations for intelligent systems, Part I: Quality criteria for definitions of intelligence. In: *Proceedings of the 10th Multidisciplinary Academic Conference*, pp. 73 - 80.

Monett, D., & Lewis, C.W. (2020b). Definitional foundations for intelligent systems, Part II: Constructing a definition and examples. In: *Proceedings of the 10th Multidisciplinary Academic Conference*, pp. 81 - 89.

Pathak, P., Pal, P.R., Shrivastava, M., & Ora, P. (2019). Fifth revolution: Applied AI & human intelligence with cyber physical systems. *International Journal of Engineering and Advanced Technology*, 8(3), pp. 23-27.

Schwab, K. (2017). The fourth industrial revolution. World Economic Forum, Currency, Geneva, pp. 1-173.

Sheppard, G. (2019). Intelligent systems: The future of intelligent systems is exciting, but making it happen will require hard work. *Research-Technology Management*, 62(4), pp. 46-52.

Wang, P. (2006). Artificial intelligence: What it is, and what it should be. In: *Proceedings of the AAAI Spring Symposium*. American Association for Artificial Intelligence, pp. 97-103.

Wilson, K.M., Helton, W.S., & Wiggins, M.W. (2013). Cognitive engineering. *Wiley Interdisciplinary Reviews: Cognitive Science*, 4(1), pp. 17-31.

Woods, D.D. (1987). Commentary: Cognitive engineering in complex and dynamic worlds. *International Journal of Man-Machine Studies*, 27(5-6), pp. 571-585.

A.1.3. References related to Chapter 2.

Abou Eddahab, F.-Z. (2020). Exemplifying smart functions for a next generation data analytics toolbox. Dissertation (TU Delft), Delft University Press, 328 p.

Ahmad, A., Paul, A., Rathore, M.M., & Chang, H. (2016). Smart cyber society: Integration of capillary devices with high usability based on cyber–physical system. *Future Generation Computer Systems*, 56, pp. 493-503.

Ansari, F., Khobreh, M., Seidenberg, U., & Sihn, W. (2018). A problem-solving ontology for human-centered cyber physical production systems. *CIRP Journal of Manufacturing Science and Technology*, 22, pp. 91-106.

Chen, H. (2017). Theoretical foundations for cyber-physical systems: A literature review. *Journal of Industrial Integration and Management*, 2(03), 1750013, pp. 1-12.

Colombo, A.W., Karnouskos, S., Kaynak, O., Shi, Y., & Yin, S. (2017). Industrial cyberphysical systems: A backbone of the fourth industrial revolution. *IEEE Industrial Electronics Magazine*, 11(1), pp. 6-16.

CPSO (2009). Cyber-physical systems - A concept map. Cyber Physical Systems Organization. http://CyberPhysicalSystems.org/CPSConceptMap.xml

CPSoS Consortium. (2015). Cyber-physical systems of systems: Research and innovation priorities initial document for public consultation. European Union's Seventh Program, No 611115. pp. 1-21.

CPSoS Consortium. (2016). Towards a European roadmap on research and innovation in engineering and management of cyber-physical systems of systems. pp. 1-62.

Deka, L., Khan, S. M., Chowdhury, M., & Ayres, N. (2018). Transportation cyber-physical system and its importance for future mobility. In: *Transportation Cyber-Physical Systems*, Elsevier, Amsterdam, pp. 1-20.

Du Bois, E. (2013). A designerly methodology for software development. Ph.D. Dissertation (TU Delft), VSSD, Delft, 252 p.

Dumitrache, I. (2010). The next generation of cyber-physical systems. *Journal of Control Engineering and Applied Informatics*, 12(2), pp. 3-4.

Falkenberg, R., Drenhaus, J., Sliwa, B., & Wietfeld, C. (2018). System-in-the-loop design space exploration for efficient communication in large-scale IoT-based warehouse systems. In: *Proceedings of the Annual IEEE International Systems Conference*, IEEE, pp. 1-6.

Gerritsen, B.H., & Horváth, I. (2012). Current drivers and obstacles of synergy in cyber-physical systems design. In: *Proceedings of the ASME 2012 International Design Engineering*

Technical Conferences. ASME. pp. 1277-1286.

Gooding, H., Lattanzio, S., Parry, G. C., & Newnes, L. (2022). Perceptions of transdisciplinary engineering: Characterisations of the transdisciplinary research approach. In: *Proceedings of the 29th International Society of Transdisciplinary Engineering Conference*, IOS Press, Amsterdam, pp. 707-716.

Hartmann, M., Viehweger, M., Desmet, W., Stolz, M., & Watzenig, D. (2017). "Pedestrian in the loop": An approach using virtual reality. In: *Proceedings of the XXVI International Conference on Information, Communication and Automation Technologies*, IEEE, pp. 1-8.

Hollnagel, E. (1993). The intelligent use of intelligent systems. In: *Intelligent Systems: Safety, Reliability and Maintainability Issue*s. Springer, Berlin, Heidelberg, pp. 42-59.

Horváth, I. (2014). What the design theory of social-cyber-physical systems must describe, explain and predict? In: *An Anthology of Theories and Models of Design*. Springer, London, pp. 99-120.

Horváth, I., & Gerritsen, B.H. (2012). Cyber-physical systems: Concepts, technologies and implementation principles. In: *Proceedings of the Ninth International Tools and Methods of Competitive Engineering Symposium – TMCE 2012*, pp. 19-36.

Horváth, I., & Tavčar, J. (2021). Designing cyber-physical systems for runtime self-adaptation: Knowing more about what we miss... *Journal of Integrated Design and Process Science*, 25(2), pp. 1-26.

Horváth. I. (2021). Connectors of smart design and smart systems. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, pp. 1-19.

Koulamas, C., & Kalogeras, A. (2018). Cyber-physical systems and digital twins in the industrial internet of things [cyber-physical systems]. *Computer*, 51(11), pp. 95-98.

Kumar, P., Mirza, K.B., Choudhury, K., Cucchiarini, M., Madry, H., & Shukla, P. (2021). Tissue regeneration through cyber-physical systems and microbots. *Advanced Functional Materials*, 31(31), 2009663.

Lee, E.A. (2007). Computing foundations and practice for cyber-physical systems: A preliminary report. University of California, Berkeley, Tech. Rep. UCB/EECS-2007-72, pp. 1-21.

Li, C. (2016). Cyber-physical solution for an engagement enhancing rehabilitation system. *Ph.D. Dissertation* (TU Delft), Delft University Press, 190 p.

Li, C., Rusák, Z., Horváth, I., Ji, L., & Hou, Y. (2014). Current status of robotic stroke rehabilitation and opportunities for a cyber-physically assisted upper limb stroke rehabilitation. In: *Proceedings of the International Tools and Methods of Competitive Engineering Symposium,* Vol. 1, pp. 899-914.

Li, Y. (2019). Utilizing dynamic context semantics in smart behavior of informing cyber-physical systems. *Ph.D. Dissertation* (TU Delft), Delft University Press, 305 p.

Lipscomb, M., Tanink. M., & Leon, J. (2023) An investigation into the development of convergence engineering. *Journal of Integrated Design and Process Science*, 26(x), pp. 1-12.

Liu, Y., Peng, Y., Wang, B., Yao, S., & Liu, Z. (2017). Review on cyber-physical systems. *IEEE/CAA Journal of Automatica Sinica*, 4(1), pp. 27-40.

Marwedel, P. (2021). Embedded system design: Embedded systems foundations of cyber-physical systems, and the internet of things. Fourth Edition. Springer Nature AG, Cham, pp. 1-446.

NRC. (2014). Convergence: Facilitating transdisciplinary integration of life sciences, physical sciences, engineering, and beyond. The National Academies Press, Washington, DC, pp. 1-x.

Pivoto, D.G., de Almeida, L.F., da Rosa Righi, R., Rodrigues, J.J., Lugli, A.B., & Alberti, A.M. (2021). Cyber-physical systems architectures for industrial internet of things applications in Industry 4.0: A literature review. *Journal of Manufacturing Systems*, 58, pp. 176-192.

Pourtalebi, S.H. (2017). System-level feature-based modeling of cyber-physical systems: A theoretical framework and methodological fundamentals. *Ph.D. Dissertation* (TU Delft), Delft University Press, 310 p.

Pourtalebi, S.H., & Horváth, I. (2016). Towards a methodology of system manifestation features-based pre-embodiment design. *Journal of Engineering Design*, 27(4-6), pp. 232-268.

Pourtalebi, S.H., Horváth, I., & Opiyo, E.Z. (2014). First steps towards a mereo-operandi theory for a system feature-based architecting of cyber-physical systems. In: *Proceedings of the Informatik 2014 Conference on Big Data*, Gesellschaft für Informatik, Stuttgart, Germany, pp. 2001-2006.

Putnik, G.D., Ferreira, L., Lopes, N., & Putnik, Z. (2019). What is a cyber-physical system? Definitions and models spectrum. *FME Transactions*, 47(4), pp. 663-674.

Quadri, I., Bagnato, A., Brosse, E., & Sadovykh, A. (2015). Modeling methodologies for cyber-physical systems: Research field study on inherent and future challenges. *Ada User Journal*, 36(4), pp. 246-253.

Rai, S., & Rai, A. (2015). Nanotechnology - The secret of fifth industrial revolution and the future of next generation. *Nusantara Bioscience*, 7(2), pp. 61-66.

Robinson, C. (2021). Cyber-physical systems from the application perspective. *HiPEAC Vision*, pp. 30-33.

Ruiz Arenas, S. (2018). Exploring the role of system operation modes in failure analysis in the context of first generation cyber-physical systems. *Ph.D. Dissertation* (TU Delft), Delft University Press, 250 p.

Sampigethaya, K., & Poovendran, R. (2013). Aviation cyber-physical systems: foundations for future aircraft and air transport. *Proceedings of the IEEE*, 101(8), pp. 1834-1855.

Schirner, G., Erdogmus, D., Chowdhury, K., & Padir, T. (2013). The future of human-in-the-loop cyber-physical systems. *Computer*, 46(1), pp. 36-45.

Shaw, I.S. (1998). What is an Intelligent System? In: *Fuzzy Control of Industrial Systems*. Springer, Boston, MA, pp. 1-3.

Stepanov, M., Musatov, V., Egorov, I., Pchelintzeva, S., & Stepanov, A. (2020). Cyber-physical control system of hardware-software complex of anthropomorphous robot: Architecture and models. In: *Cyber-Physical Systems: Advances in Design & Modelling*, Springer, Cham, pp. 13-24.

Suh, S.C., Tanik, U.J., Carbone, J.N., & Eroglu, A. (Eds.). (2014). Applied cyber-physical systems. Springer, New York.

Talsania, H.N., Raval, T.F., & Chaudhary, K.A. (2017). A survey paper on cyber physical systems. *International Journal for Scientific Research & Development*, 5(8), pp. 13-16.

Tan, Y., Goddard, S., & Perez, L. C. (2008). A prototype architecture for cyber-physical systems. *ACM Sigbed Review*, 5(1), pp. 1-2.

Tepjit, S. (2022). Designing an active recommender framework to support the development of reasoning mechanisms for smart cyber-physical systems. *Ph.D. Dissertation* (TU Delft), Delft University Press, 317 p.

Togay, C. (2014). Chapter 8: Axiomatic design theory for cyber-physical system. In: *Applied Cyber-Physical Systems*, Springer Science + Business Media, New York, N.Y., pp. 85-100.

van Doorn, E. (2018). Information engineering for supporting situation awareness of nautical traffic management operators. *Ph.D. Dissertation* (TU Delft), Delft University Press, 280 p.

Vegetti, M.M., Böhm, A., Leone, H.L., & Henning, G.P. (2021). SCONTO: A modular ontology for supply chain representation. In: *Domain Ontologies for Research Data Management in Industry Commons of Materials and Manufacturing*, 40. pp. 40-55.

Wan, J., Yan, H., Suo, H., & Li, F. (2011). Advances in cyber-physical systems research. *KSII Transactions on Internet and Information Systems*, 5(11), pp. 1891-1908.

Wetter, K.J. (2006). Implications of technologies converging at the nano-scale. *Nanotechnologien nachhaltig gestalten. Konzepte und Praxis fur eine verantwortliche Entwicklung und Anwendung.* Institut für Kirche und Gesellschaft, Iserlohn, Germany, pp. 31-41.

Weyns, D., Andersson, J., Caporuscio, M., Flammini, F., Kerren, A. & Lowe, W. (2021). A research agenda for smarter cyber-physical systems. *Journal of Integrated Design and Process Science*, 25(2), pp. 27-47.

Willems, J.C. (1991). Paradigms and puzzles in the theory of dynamical systems. IEEE *Transactions on Automatic Control*, 36(3), pp. 259-294.

Xu, H., Yu, W., Griffith, D., & Golmie, N. (2018). A survey on industrial Internet of Things: A cyber-physical systems perspective. *IEEE Access*, 6, pp. 78238-78259.

Xue, Y., Rodriguez, S., & Bogdan, P. (2016). A spatio-temporal fractal model for a CPS approach to brain-machine-body interfaces. In: *Proceedings of the Design, Automation & Test in Europe Conference & Exhibition – DATE '16*, IEEE, pp. 642-647.

Yamanobe, N., Wan, W., Ramirez-Alpizar, I.G., Petit, D., Tsuji, T., Akizuki, S., Hashimoto, M., Nagata, K., & Harada, K. (2017). A brief review of affordance in robotic manipulation research. *Advanced Robotics*, 31(19-20), pp. 1086-1101.

Zhou, Y., Yu, F.R., Chen, J., & Kuo, Y. (2019). Cyber-physical-social systems: A state-of-the-art survey, challenges and opportunities. *IEEE Communications Surveys & Tutorials*, 22(1), pp. 389-425.

A1.4. References related to Chapter 3.

Abbod, M.F., Linkens, D.A., Mahfouf, M., & Dounias, G. (2002). Survey on the use of smart and adaptive engineering systems in medicine. *Artificial Intelligence in Medicine*. 26(3), pp. 179-209.

Andry, P., Gaussier, P., Moga, S., Banquet, J.P., & Nadel, J. (2001). Learning and communication via imitation: An autonomous robot perspective. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 31(5), pp. 431-442.

Black, W.S., Haghi, P., & Ariyur, K.B. (2014). Adaptive systems: History, techniques, problems, and perspectives. *Systems*, 2(4), pp. 606-660.

Borches, P.D., & Bonnema, G.M. (2008). On the origin of evolvable systems: Evolvability or extinction. In: *Proceedings of the International TMCE 2008 Symposium*, Vol. 2., April 21–25, 2008, Kusadasi, Turkey, TU Delft, pp. 1351-1352.

Brown, C.A. (2006). The application of complex adaptive systems theory to clinical practice in rehabilitation. *Disability and Rehabilitation*, 28(9), pp. 587-593.

Brundage, M. (2016). Artificial intelligence and responsible innovation. In: *Fundamental Issues of Artificial Intelligence*, Springer International Publishing Switzerland, pp. 543-554.

Chaterji, S., Naghizadeh, P., Alam, M. A., Bagchi, S., Chiang, M., Corman, D., ... & Weller, J. (2019). Resilient cyberphysical systems and their application drivers: A technology roadmap. arXiv preprint arXiv:2001.00090.

Claudi, A., Sernani, P., Dolcini, G., Palazzo, L., & Dragoni, A. F. (2013). A hierarchical hybrid model for intelligent cyber-physical systems. In: *Proceedings of the 11th Workshop on Intelligent Solutions in Embedded Systems*, IEEE, pp. 1-6.

Demetriou, A., Makris, N., Kazi, S., Spanoudis, G., & Shayer, M. (2018). The developmental trinity of mind: Cognizance, executive control, and reasoning. *Wiley Interdisciplinary Reviews: Cognitive Science*, 9(4), e1461.

Endsley, M.R. (2018). Automation and situation awareness. In: *Automation and Human Performance: Theory and Applications*, CRC Press, pp. 163-181.

Erickson, R.P. (2014). Are humans the most intelligent species? *Journal of Intelligence*, 2(3), pp. 119-121.

Frei, C.W., Kraus, F.J., & Blanke, M. (1999). Recoverability viewed as a system property. In: *Proceedings of the 1999 European Control Conference*, IEEE, pp. 2197-2202.

Gabor, T., Belzner, L., Kiermeier, M., Beck, M. T., & Neitz, A. (2016). A simulation-based architecture for smart cyber-physical systems. In: *Proceedings of the 2016 International*

Conference on Autonomic Computing, IEEE, pp. 374-379.

Greenberg S., Gutwin, C., & Cockburn, A. (1996). Awareness through fisheye views in relaxed-WYSIWIS groupware. In: *Proceedings of the Graphics Interface Conference,* Morgan-Kaufmann, pp. 28-38.

Haböck, U., Klier, J., Schwenninger, J., & Maier, S. (2016). System adaptation as key technology towards automated driving. *ATZ Worldwide*, 118(4), pp. 26-31.

Higuchi T. (1993). Evolvable hardware: A first step towards building a Darwin machine. In: *Proceedings of the 2nd International Conference on Simulated Behaviour*, MIT Press, pp. 417-424.

Hou, Z.S., & Wang, Z. (2013). From model-based control to data-driven control: Survey, classification and perspective. *Information Sciences*, 235, pp. 3-35.

Hunt, E., & Jaeggi, S.M. (2013). Challenges for research on intelligence: Commentary. *Journal of Intelligence*, 1(1), pp. 36-54.

Leitão, P., Colombo, A.W., & Karnouskos, S. (2016). Industrial automation based on cyber-physical systems technologies: Prototype implementations and challenges. *Computers in Industry*, 81, pp. 11-25.

Li, C., Rusák, Z., Horváth, I., & Ji, L. (2016). Validation of the reasoning of an entry-level cyber-physical stroke rehabilitation system equipped with engagement enhancing capabilities. *Engineering Applications of Artificial Intelligence*, 56, pp. 185-199.

Mylonas, P., Spyrou, E., Avrithis, Y., & Kollias, S. (2009). Using visual context and region semantics for high-level concept detection. *IEEE Transactions on Multimedia*, 11(2), pp. 229-243.

NASEM. (2016). A 21st century cyber-physical systems education. National Academies of Sciences, Engineering, and Medicine. Washington, DC: The National Academies Press. doi:10.17226/23686.

Ning, H., Liu, H., Ma, J., Yang, L.T., & Huang, R. (2016). Cybermatics: Cyber–physical-social-thinking hyperspace based science and technology. *Future Generation Computer Systems*, 56, pp. 504-522.

Oreizy, P., Gorlick, M.M., Taylor, R.N., Heimbigner, D., Johnson, G., Medvidovic, N., ... & Wolf, A.L. (1999). An architecture-based approach to self-adaptive software. *IEEE Intelligent Systems*, 14(3), pp. 54-62.

Reggia, J.A. (2013). The rise of machine consciousness: Studying consciousness with computational models. *Neural Networks*, 44, pp. 112-131.

Signorelli, C.A. (2017). Types of cognition and its implications for future high-level cognitive machines. In: *Proceedings of the AAAI Spring Symposium Series*, Technical Report SS-17-07, AAAI, pp. 622-627.

Tackett, M.W., Mattson, C.A., & Ferguson, S.M. (2014). A model for quantifying system evolvability based on excess and capacity. *Journal of Mechanical Design*, 136(5), pp. 051002

Thekkilakattil, A., & Dodig-Crnkovic, G. (2015). Ethics aspects of embedded and cyber-physical systems. In: *Proceedings of the 39th Annual Computer Software and Applications Conference*, Vol. 2, IEEE, pp. 39-44.

Tononi, G. (2012). Integrated information theory of consciousness: An updated account. *Archives Italiennes De Biologie*, 150(2-3), pp. 56-90.

Tyagi, A.K., Fernandez, T.F., Mishra, S., & Kumari, S. (2021). Intelligent automation systems at the core of industry 4.0. In: *Proceedings of the 20th International Conference on Intelligent Systems Design and Applications*, Springer International Publishing, Cham, pp. 1-18.

Weyns, D., Bäck, T., Vidal, R., Yao, X., & Belbachir, A. N. (2022). The vision of self-evolving computing systems. *Journal of Integrated Design and Process Science*, (Preprint), pp. 1-17.

Wu, L., & Kaiser, G. (2012). An autonomic reliability improvement system for cyber-physical systems. In: *Proceedings of the 14th International Symposium on High-Assurance Systems Engineering*, IEEE, pp. 56-61.

A1.5. References related to Chapter 4.

Agresto, J. (1983). The humanities and the social sciences. *JSTOR – PS 16*, pp. 543-545.

Aikins, J.S. (1983). Prototypical knowledge for expert systems. *Artificial Intelligence*, 20(2), pp. 163-210.

Boas, G. (1957). The problem of the humanities. *The Journal of General Education,* 10, pp. 205–216.

Bohman, J. (2002). How to make a social science practical? Pragmatism, critical social science and multiperspectival theory. *Millennium*, 31(3), pp. 499-524.

Boschert S, Heinrich C and Rosen R (2018) Next generation digital twin. In: *Proceedings of the Twelfth International Symposium of the Tools and Methods of Competitive Engineering*, Las Palmas de Gran Canaria, Spain, pp. 209–2017.

Broderick, F. (1983). What are the humanities? The Educational Resources Information Center, pp. 1–3.

Canton, J. (2004). Designing the future: NBIC technologies and human performance enhancement. *Annals-New York Academy of Sciences*, 1013, pp. 186-198.

Cross, N. (2001). Designerly ways of knowing: Design discipline versus design science (The 1920s and the 1960s, two important periods in the modern history of design), *Design Issues*, 17(3), pp. 49-55.

Džeroski S, Langley P and Todorovski L (2007) Computational discovery of scientific knowledge. In: *Computational Discovery of Scientific Knowledge: Introduction, Techniques, and Applications in Environmental and Life Sciences*. Springer, Berlin, Heidelberg, New York, pp. 1–14.

Eekels, J., & Roozenburg, N. (1991). A methodological comparison of the structures of scientific research and engineering design: Their similarities and differences. *Design Studies*, 12(4), pp. 197-203.

Epstein, Z., Payne, B.H., Shen, J.H., Dubey, A., Felbo, B., Groh, M., ... & Rahwan, I. (2018). Closing the AI knowledge gap. arXiv preprint arXiv:1803.07233.

Frayling, C. (1993). Research in art and design. Royal College of Art, London. 1(1), pp. 1-5.

French, S. (2007). Science: Key concepts in philosophy. Continuum Books, London, pp. 43–71.

Friedman, K. (2017). Editorial: Conversation, discourse, and knowledge. *She Ji - The Journal of Design, Economics, and Innovation*, 3(2), pp. 75-81.

Gedenryd, H. (1998). How designers work: Making sense of authentic cognitive activities (Vol. 75). Lund University, pp. 1-123

Giard, J., & Gilles, W. (2001). Delta knowledge: Its place and significance in industrial design education. In: *Proceedings of the ICSID Educational Seminar*, pp. 165-169.

Gilles, W., & Paquet, G. (1989). On delta knowledge. In: *Edging Toward the Year 2000*. Canadian Federation of Deans of Management and Administrative Studies, Ottawa, pp. 15–30

Gordon, E.C. (2012). Is there propositional understanding? *Logos & Episteme,* 3, pp. 181–192.

Graham, P.J., & Dickinson, H.D. (2007). Knowledge-system theory in society: Charting the growth of knowledge-system models over a decade, 1994-2003, *Journal of the American Society for Information Science and Technology*, 58(14), pp.2372–2381.

Horváth, I., & Duhovnik, J. (2005). Towards a better understanding of the methodological characteristics of engineering design research. In: *Proceedings of the International Design Engineering Technical Conferences and Computers and Information in Engineering Conference* (Vol. 47403), ASME, pp. 905-919.

Jashapara, A. (2010). Knowledge management: An integrated approach. Pearson Education Limited, London, pp. 1-376.

Lakatos, I. (1970). History of science and its rational reconstructions. In: *Proceedings of the Biennial Meeting of the Philosophy of Science Association*, Vol. 1970, PSA, D. Reidel Publishing, pp. 91–136.

Levi, A.W. (1983). The humanities: their essence, nature, future. *Journal of Aesthetic Education,* 17, pp. 5–17.

Liu, S.C., Tomizuka, M., & Ulsoy, A.G. (2004). Challenges and opportunities in the engineering of intelligent systems. In: *Proceedings of the 4th International Workshop on Structural Control*, New York, pp. 295-300.

Machlup, F. (1982). Knowledge: Its creation, distribution and economic significance. Vol. II. *The Branches of Learning,* Princeton University Press, Princeton, NJ, pp. 1–207.

Machlup, F. (1980). Knowledge: Its creation, distribution and economic significance. Vol. I. *Knowledge and Knowledge Production,* Princeton University Press, Princeton, NJ, pp. 1–245.

Machlup, F. (1984). Knowledge: ts creation, distribution and economic significance. Vol. III. *The Economics of Information and Human Capital*, Princeton University Press, Princeton, NJ: pp. 1–611.

Mayntz, R. (1990). The influence of natural science theories on contemporary social science. Discussion Paper 90/7. Max-Planck-Institut fur Gesellschaftsforschung, pp. 1-65.

Mendelsohn, E. (1977). The social construction of scientific knowledge. In: *The Social Production of Scientific Knowledge*. Springer, Dordrecht, pp. 3–26.

Mises, L. (1942). Social science and natural science. *Journal of Social Philosophy & Jurisprudence*, 7(3), pp. 240-253.

Myers, F.M. (1967). Toward a theory of the humanities. *New Mexico Quarterly*, 37, 2-x.

Narváez, L.M. (2000). Design's own knowledge, *Design Issues*, 16(1), pp. 36-51.

Popper, K. (1962). Conjectures and refutations: The growth of scientific knowledge. Basic Books Publishers, New York, London.

Radermacher, F.J. (1996). Cognition in systems. *Cybernetics and Systems*, 27(1), pp. 1-41.

Reber, R. & Bullot, N.J. (2019). Conditional objectivism: A strategy for connecting the social sciences and practical decision-making. Springer Nature AG, Switzerland, pp. 73-92.

Rittel, H.W., & Webber M.M. (1973), Dilemmas in a general theory of planning. *Policy Sciences*, 4(2), pp.155–169.

Rosenberg, A. (2000). Philosophy of science: A contemporary introduction. Routledge, London and New York, pp. 96–102.

Schön, D.A. (1988). From technical rationality to reflection-in-action. In: *Professional judgment: A reader in clinical decision making*, Cambridge University Press, London, pp. 60-77.

Sheng, J., Amankwah-Amoah, J., & Wang, X. (2019). Technology in the 21st century: New challenges and opportunities. *Technological Forecasting and Social Change*, 143, pp. 321-335.

Smith, R. (2015) Thinking and making: art and craft in library of congress classification. *Library Philosophy and Practice. Online journal*, Paper number: 1299, pp. 1–14.

Stehr, N. (1996). The fragility of knowledge societies. In: *Proceedings of the Workshop on High-tech and Sustainability*, September 3, 1996, Bratislava, Slovak Republic, pp. 1-14.

Sumari, A.D., & Ahmad, A.S. (2017). Knowledge-growing system: The origin of the cognitive artificial intelligence. In: *Proceedings of the 6th International Conference on Electrical Engineering and Informatics*, IEEE, pp. 1-7.

Tang, S. (2011). Foundational paradigms of social sciences. *Philosophy of the Social Sciences*, 41(2), pp. 211-249.

Tebes, J.K. (2017). Foundations for a philosophy of science of community psychology: Perspectivism, pragmatism, feminism, and critical theory. In: *APA Handbook of Community Psychology*, American Psychological Association, pp. 21-40.

Uluoğlu, B. (2000). Design knowledge communicated in studio critiques. *Design Studies*, 21(1), pp.

33-58.

Valsiner, J. (2019). Theory and history in the human and social sciences. In: *Social Philosophy of Science for the Social Sciences*, Springer, New York, NY, pp. x-x.

Van Aken, J.E. (2005) Valid knowledge for the professional design of large and complex design processes. *Design Studies*, 26(4), pp. 379-404.

Vaziri, H., Tay, L., Keith, M.G., & Pawelski, J.O. (2019). History, literature, and philosophy: A systematic review of positive functioning. *The Journal of Positive Psychology,* 14, pp. 695–723.

Wayne, J.J. (2013). Physics laws of social science. *Munich Personal RePEc Archive*, Paper No. 47811, pp. 1-11.

Wray, K.B. (2007). Who has scientific knowledge? *Social Epistemology*. 21, pp. 337–347.

Ziori, E. (2004). Implict knowledge representation. In: Not identified publication, pp. 217-231.

A1.6. References related to Chapter 5.

Aamodt, A. (1995). Knowledge acquisition and learning by experience - The role of case-specific knowledge, In: *Machine Learning and Knowledge Acquisition - Integrated Approaches*, Ch. 8, Academic Press, pp. 197-245.

Aussenac-Gilles, N., Charlet, J., & Reynaud, C. (2020). Knowledge engineering. In: *A Guided Tour of Artificial Intelligence Research*, Springer, Cham, pp. 733-768.

Baumgarten, A.G. (1986). Aesthetica – 1750. Hildesheim, Olms, pp. 1-x.

Boulding, K.E. (1987). The epistemology of complex systems. *European Journal of Operation Research.*, 30, pp. 110–116.

Bratianu, C., & Andriessen, D. (2008). Knowledge as energy: A metaphorical analysis. In: *Proceedings of the 9th European Conference on Knowledge Management*, Reading, Southampton, UK, 4–5 September 2008; pp. 75–82.

Ferrier, J.F. (1866). Lectures on Greek philosophy and other philosophical remains of James Frederick Ferrier; William Blackwood and Sons: Edinburgh, London.

Figueiredo, A.D. (2008). Toward an epistemology of engineering. In: Proceedings of the Workshop on Philosophy and Engineering, London, UK, 10–12 November 2008, pp. 1-x.

Goldman, A.I. (1986). Epistemology and cognition; Harvard University Press: Cambridge, MA.

Helmer, O., & Rescher, N. (1959). On the epistemology of the inexact sciences. *Management Science*, 6, pp. 25–52.

Hjørland, B. (2009). Concept theory. *Journal of the American Society for Information Science and Technology*, 60, pp. 1519–1536.

Holsapple, C.W. (2004). Knowledge and its attributes. In: *Handbook on Knowledge Management*, Springer, Berlin, Heidelberg, pp. 165-188.

Hooker, C. (2011). Introduction to philosophy of complex systems: Part A: Towards a framework for complex systems. In *Philosophy of Complex Systems*; Elsevier: Amsterdam, The Netherlands, pp. 3–90.

Li, D. (2013). An ontological perspective of value. Part II - Front Matter. In: *Value Theory: A Research into Subjectivity*. Springer Science & Business Media: Berlin/Heidelberg, pp. 99–101.

Möbus, C. (1996). Towards an epistemology on intelligent design and modelling environments: the hypothesis testing approach. In: *Proceedings of the European Conference on Artificial Intelligence in Education*, Lisbon, Portugal, 30 September–2 October 1996, pp. 52–58.

Mohaghegh, D.M., & McCauley, M. (2016). Computational thinking: The skill set of the 21st century. *International Journal of Computer Science and Information Technologies*, 7(3), 1524-1530.

Nguyen, K.V. (2015). Gnoseology: In: *Relation to Truth, Knowledge and Metaphysics*. PhilPapers, http://philpapers.org/rec/NGUGIR.

Nikitchenko, M. (2011) Gnoseology-based approach to foundations of informatics. In: *Proceedings* *of the 7th International Conference on ICT in Education, Research and Industrial Applications: Integration, Harmonization and Knowledge Transfer*, Kherson, Ukraine, 4–7 May 2011; p. 1-27

Ratcliff, R. (2013). Applying epistemology to system engineering: An illustration. *Procedia Computer Science*, 16, pp. 393–402.

Solaz-Portolés, J.J., & Sanjosé, V. (2008). Types of knowledge and their relations to problem solving in science: Directions for practice, *Sísifo: Educational Sciences Journal*, 6, pp. 105–112.

Stenning, K., & Van Lambalgen, M. (2012). Human reasoning and cognitive science. MIT Press, Boston, pp. 1-422.

Stokes, D.E. (1997). Pasteur's quadrant: Basic science and technological innovation; Brookings Institution Press: Washington, DC.

Tani, J. (2014). Self-organization and compositionality in cognitive brains: A neurorobotics study. *Proceedings of the IEEE*, 102(4), pp. 586-605.

Thannhuber, M.J. (2005). Theory of system knowledge: An introduction to the proposed theoretical framework - Understanding and describing knowledge and intelligence on an abstract system level. In: *The Intelligent Enterprise: Theoretical Concepts and Practical Implications*; Springer: Heidelberg, Germany, pp. 53–66.

A1.7. References related to Chapter 6.

Albright, J., & Runehov, A. (2013). Reductionism. In: *Encyclopedia of Sciences and Religions*, Springer, New York, NY, pp. 1006-1008.

Baker, V.R. (2007). Flood hazard science, policy, and values: A pragmatist stance. *Technology in Society*, *29*(2), pp. 161-168.

Behl, D.V., & Ferreira, S. (2014). Systems thinking: An analysis of key factors and relationships. *Procedia Computer Science*, 36, pp. 104-109.

Braun, W. (2002). The system archetypes. *Systems*, 2002, pp. 27-x.

Buckle Henning, P., & Chen, W. (2012). Systems thinking: Common ground or untapped territory? *Systems Research and Behavioral Science*, 29(5), pp. 470-483.

Burge, S. (2015). An overview of the soft systems methodology. *System Thinking: Approaches and Methodologies,* pp. 1-14.

Cabrera, D., Cabrera, L., & Powers, E. (2015). A unifying theory of systems thinking with psychosocial applications. *Systems Research and Behavioral Science*, 32(5), pp. 534-545.

Checkland, P. (2000). Systems thinking, systems practice: includes a 30-year retrospective. *Journal of the Operational Research Society*, 51(5), pp. 647-647.

Daellenbach, H.G., McNickle, D.C., & Dye, S. (2012). Management science: Decision making through systems thinking (2nd edition). Palgrave Macmillan, Basingstoke, pp. 1-598.

Dorani, K., Mortazavi, A., Dehdarian, M., Mahmoudi, H., Khandan, M., & Mashayekhi, A. (2015). Developing question sets to assess systems thinking skills, In: *Proceedings of the 33rd International Conference of the System Dynamics Society*, Cambridge, MA, July 19-23, 2015, pp. 1-17.

Drack, M., & Apfalter, W. (2007). Is Paul A. Weiss' and Ludwig von Bertalanffy's system thinking still valid today? *Systems Research and Behavioral Science: The Official Journal of the International Federation for Systems Research*, 24(5), pp. 537-546.

Fazey, I., Schäpke, N., Caniglia, G., Hodgson, A., Kendrick, I., Lyon, C., ... & Saha, P. (2020). Transforming knowledge systems for life on Earth: Visions of future systems and how to get there. *Energy Research & Social Science*, 70, 101724.

Forrester, J.W. (1990). Principles of systems. Pegasus Communications: Waltham, MA.

Gradinarov, B. (2015). Social sciences in front of the paradigm of systems thinking. *Humanities and Social Sciences*, 2(1-2), pp. 61-72.

Hammond, D. (2005). Philosophical and ethical foundations of systems thinking. *TripleC:*

Communication, Capitalism & Critique. Open Access Journal for a Global Sustainable Information Society, 3(2), pp. 20-27.

Hürlimann, M. (2009). System thinking. In: *Dealing with Real-World Complexity*. Gabler, pp. 59-78.

Jackson, M.C. (2009). Fifty years of systems thinking for management. *Journal of the Operational Research Society*, 60(1), pp. S24-S32.

Jaradat, R.M., & Keating, C.B. (2016). Systems thinking capacity: implications and challenges for complex system governance development. *International Journal of System of Systems Engineering*, 7(1-3), pp. 75-94.

Keating, C.B., & Gheorghe, A.V. (2016). Systems thinking: Foundations for enhancing system of systems engineering. In: *Proceedings of the 11th System of Systems Engineering Conference*, IEEE, pp. 1-6.

Kim, D.H., & Anderson, V. (1998). Systems archetype basics. Pegasus Communications Inc., Waltham, MA., pp. 1-150.

Mahmoudi, H., Dorani, K., Dehdarian, A., Khandan, M., & Mashayekhi, A.N. (2019). Does systems thinking assessment demand a revised definition of systems thinking. In: *Proceedings of the 37th International Conference of the System Dynamics Society*. Albuquerque, NM, pp. 1-17

Monat, J.P., & Gannon, T.F. (2015). What is systems thinking? A review of selected literature plus recommendations. *American Journal of Systems Science*, 4(1), pp. 11-26.

Oh, J., Meneguzzi, F., Sycara, K., & Norman, T.J. (2013). Prognostic normative reasoning. *Engineering Applications of Artificial Intelligence*, 26(2), pp. 863-872.

Pennock, M.J., & Wade, J.P. (2015). The top 10 illusions of systems engineering: A research agenda. *Procedia Computer Science*, 44, pp. 147-154.

Pickel, A. (2007). Rethinking systems theory: A programmatic introduction. *Philosophy of the Social Sciences*, 37(4), pp. 391-407.

Richmond, B. (1991). Systems thinking: Four key questions. High Performance Systems. Inc., Lebanon, NH, USA, pp. 1-x.

Richmond, B. (1993). Systems thinking: Critical thinking skills for the 1990s and beyond. *System Dynamics Review*, 9(2), pp. 113–133.

Senge, P.M. (2006). The fifth discipline: The art and practice of the learning organization, *Doubleday/Currency*, New York, pp. 1-436.

Stave, K., & Hopper, M. (2007). What constitutes systems thinking? A proposed taxonomy. In: *Proceedings of the 25th International Conference of the System Dynamics Society*. July 29-August 3, 2007, Boston, MA, pp. 1-24.

Steels, L., & López de Mantaras, R. (2018). The Barcelona declaration for the proper development and usage of artificial intelligence in Europe. *AI Communications*, 31(6), pp. 485-494.

Straus, J. (2021). Artificial intelligence – Challenges and chances for Europe. *European Review*, 29(1), pp. 142-158.

Sun, D., Hyland, P., & Cui, H. (2014). A designed framework for delivering systems thinking skills to small business managers. *Systems*, 2(3), pp. 297-312.

Tepjit, S., Horváth, I., & Rusák, Z. (2019). The state of framework development for implementing reasoning mechanisms in smart cyber-physical systems: A literature review. *Journal of Computational Design and Engineering*, 6(4), pp. 527-541.

Ulrich, W. (2013). Critical systems thinking. In: *Encyclopedia of Operations Research and Management Science*, 2, pp. 314-326.

Waddock, S. (2016). Foundational memes for a new narrative about the role of business in society. *Humanistic Management Journal*, 1(1), pp. 91-105.

Wan, P.Y. (2011). Emergence à la systems theory: Epistemological Totalausschluss or ontological novelty? *Philosophy of the Social Sciences*, 41(2), pp. 178-210.

Wendt, C., Frisina, L., & Rothgang, H. (2009). Healthcare system types: A conceptual framework for comparison. *Social Policy & Administration*, 43(1), pp. 70-90.

Zexian, Y., & Xuhui, Y. (2010). A revolution in the field of systems thinking - A review of Checkland's system thinking. *Systems Research and Behavioral Science: The Official Journal of the International Federation for Systems Research*, 27(2), pp. 140-155.

A1.8. References related to Chapter 7.

AbdulRahman, S., Tout, H., Ould-Slimane, H., Mourad, A., Talhi, C., & Guizani, M. (2020). A survey on federated learning: The journey from centralized to distributed on-site learning and beyond. *IEEE Internet of Things Journal*, 8(7), pp. 5476-5497.

Abu-Salih, B. (2021). Domain-specific knowledge graphs: A survey. *Journal of Network and Computer Applications,* 185, 103076.

Allen, J., Blaylock, N., & Ferguson, G. (2002). A problem solving model for collaborative agents. In: *Proceedings of the first International Joint Conference on Autonomous Agents and Multiagent Systems: Part 2*, pp. 774-781.

Amin, S., Usman, M., Sohail, N., & Aslam, S. (2018). Relationship between intellectual capital and financial performance: The moderating role of knowledge assets. *Pakistan Journal of Commerce and Social Sciences*, 12(2), pp. 521-547.

Athanasiadis, I.N. (2005). A review of agent-based systems applied in environmental informatics. In: *Proceedings of the Modeling and Simulation - MODSIM - Workshop*, pp. 1574-1580.

Bharat, K.A., & Cardelli, L. (1995). Migratory applications. In: *Proceedings of the 8th Annual ACM Symposium on User Interface and Software Technology*, pp. 132-142.

Cardoso, R.C., & Ferrando, A. (2021). A review of agent-based programming for multi-agent systems. *Computers,* 10(2), pp. 1-15.

Chandrasekaran, B. (1986). Generic tasks in knowledge-based reasoning: High-level building blocks for expert system design. *IEEE Intelligent Systems*, 1(03), 23-30.

Collins, C.J., & Smith, K.G. (2006). Knowledge exchange and combination: The role of human resource practices in the performance of high-technology firms. *Academy of Management Journal*, 49(3), pp. 544-560.

Gunasekaran, S.S., Mostafa, S.A., & Ahmad, M. (2015). Knowledge transfer model in collective intelligence theory. In: *Advances in Intelligent Informatics*. Springer, Cham, pp. 481-491.

Hogan, A., Blomqvist, E., Cochez, M., d'Amato, C., Melo, G. D., Gutierrez, C., ... & Zimmermann, A. (2021). Knowledge graphs. *ACM Computing Surveys*, 54(4), pp. 1-37.

Hur, A., Janjua, N., & Ahmed, M. (2021). A survey on state-of-the-art techniques for knowledge graphs construction and challenges ahead. In: *Proceedings of the Fourth International Conference on Artificial Intelligence and Knowledge Engineering*, IEEE, pp. 99-103.

Iglesias, C.A., Garijo, M., & González, J.C. (1998). A survey of agent-oriented methodologies. In: *Proceedings of the International Workshop on Agent Theories, Architectures, and Languages*. Springer, Berlin, Heidelberg, pp. 317-330

Insa-Cabrera, J., & Hernández-Orallo, J. (2013). Interaction settings for measuring (social) intelligence in multi-agent systems. In: *Proceedings of the ReteCog Interaction Workshop*, pp. 1-4.

Ji, S., Pan, S., Cambria, E., Marttinen, P., & Philip, S.Y. (2021). A survey on knowledge graphs: Representation, acquisition, and applications. *IEEE Transactions on Neural Networks and Learning Systems*, 33(2), pp. 494-514.

Khan, L.U., Saad, W., Han, Z., & Hong, C.S. (2021). Dispersed federated learning: Vision, taxonomy, and future directions. *IEEE Wireless Communications*, 28(5), 192-198.

Leitão, P., Karnouskos, S., Ribeiro, L., Lee, J., Strasser, T., & Colombo, A.W. (2016). Smart agents in industrial cyber–physical systems. *Proceedings of the IEEE*, 104(5), pp. 1086-1101.

Leitão, P., Queiroz, J., & Sakurada, L. (2022). Collective intelligence in self-organized industrial cyber-physical systems. *Electronics*, 11(19), 3213, pp. 1-21.

Liu, A., Zhang, D., Wang, Y., & Xu, X. (2022). Knowledge graph with machine learning for product design. *CIRP Annals - Manufacturing Technology*, 71, pp. 117-120.

Lykourentzou, I., Vergados, D.J., Kapetanios, E., & Loumos, V. (2011). Collective intelligence systems: Classification and modeling. *Journal of Emerging Technologies in Web Intelligence*, 3(3), pp. 217-226.

McDermott, J. 1990. Developing software is like talking to Eskimos about snow. In: *Proceedings of the Eighth National Conference on Artificial Intelligence*, American Association for Artificial Intelligence, Menlo Park, CA. pp. 1130–1133.

McMahan, H.B., Yu,F. X., Richtarik, P., Suresh, A.T., & Bacon, D. (2016). Federated learning: Strategies for improving communication efficiency. In: *Proceedings of the 29th Conference on Neural Information Processing Systems*, Barcelona, Spain, pp. 5-10.

Musil, J., Musil, A., & Biffl, S. (2015). Introduction and challenges of environment architectures for collective intelligence systems. In: *Agent Environments for Multi-Agent Systems IV*, Springer, Cham, pp. 76-94.

Neches, R., Fikes, R.E., Finin, T., Gruber, T., Patil, R., Senator, T., & Swartout, W.R. (1991). Enabling technology for knowledge sharing. *AI Magazine*, 12(3), pp. 36-36.

Neyshabur, B., Sedghi, H., & Zhang, C. (2020). What is being transferred in transfer learning? *Advances in Neural Information Processing Systems*, 33, pp. 512-523.

Noy, N., Gao, Y., Jain, A., Narayanan, A., Patterson, A., & Taylor, J. (2019). Industry-scale Knowledge Graphs: Lessons and Challenges: Five diverse technology companies show how it's done. Queue, 17(2), 48-75.

O'Neill, V., & Soh, B. (2022). Improving fault tolerance and reliability of heterogeneous multi-agent IoT systems using intelligence transfer. *Electronics*, 11(17), 2724, pp. 1-38.

O'Leary, D.E. (1998). Enterprise knowledge management. *Computer*, 31(3), pp. 54-61.

Pan, S., & Yang, Q. (2010). A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10), pp. 1345–1359.

Pratt, M.J. (2005). ISO 10303, the STEP standard for product data exchange, and its PLM capabilities. *International Journal of Product Lifecycle Management*, 1(1), pp. 86-94.

Rudowsky, I. (2004). Intelligent agents. *Communications of the Association for Information Systems*, 14(1), pp. 275-290.

Sarcheshmehpour, Y., Leinonen, M., & Jung, A. (2021). Federated learning from big data over networks. In: *Proceedings of the International Conference on Acoustics, Speech and Signal Processing*, IEEE, pp. 3055-3059.

Sattler, F., Müller, K., & Samek, W. (2020). Clustered federated learning: Model-agnostic distributed multitask optimization under privacy constraints. In: *IEEE Transactions on Neural Networks and Learning Systems*, IEEE, New York, pp. 1-x.

Schneider, E.W. (1973). Course modularization applied: The interface system and its implications for sequence control and data analysis. *Report No: HumBRO-PP-10-73*, Human Resources Research Organization, Alexandria, VA, pp.1-22.

Shaheen, M., Farooq, M.S., Umer, T., & Kim, B.S. (2022). Applications of federated learning: Taxonomy, challenges, and research trends. *Electronics*, 11(4), 670.

Smirnov, A.V., Levashova, T., Shilov, N., & Ponomarev, A. (2019). Multi-aspect ontology for interoperability in human-machine collective intelligence systems for decision support. In: *Proceedings of KEOD*, pp. 458-465.

Smith, H.N., & Poulter, K.J. (1993). The elements of an open KBS infrastructure. *Future Generation Computer Systems*, 9(4), pp. 349-369.

Sulis, W. (1997). Fundamental concepts of collective intelligence. *Nonlinear Dynamics, Psychology, and Life Sciences*, 1(1), pp. 35-53.

Sycara, K., Pannu, A., Willamson, M., Zeng, D., & Decker, K. (1996). Distributed intelligent agents. *IEEE Expert*, 11(6), pp. 36-46.

Tian, L., Zhou, X., Wu, Y. P., Zhou, W. T., Zhang, J. H., & Zhang, T. S. (2022). Knowledge graph and knowledge reasoning: A systematic review. *Journal of Electronic Science and Technology*, 100159, pp. 1-x.

Tiwari, S., Al-Aswadi, F. N., & Gaurav, D. (2021). Recent trends in knowledge graphs: Theory and practice. *Soft Computing*, 25(13), 8337-8355.

Torrey, L., & Shavlik, J. (2010). Transfer learning. In: *Handbook of Research on Machine Learning Applications and Trends: Algorithms, Methods, and Techniques*. IGI Global, pp. 242-264.

Vial, G. (2021). Understanding digital transformation: A review and a research agenda. *Managing Digital Transformation*, pp. 13-66.

Weiss, K., Khoshgoftaar, T.M., & Wang, D. (2016). A survey of transfer learning. *Journal of Big data*, 3(1), pp. 1-40.

Williams, A.B. (2004). Learning to share meaning in a multi-agent system. *Autonomous Agents and Multi-Agent Systems*, 8(2), pp. 165-193.

Williams, A.E. (2021). Approximating an artificial general intelligence or a general collective intelligence. *International Journal of Collaborative Intelligence*, 2(3), pp. 210-223.

Xu, Y., & Qi, H. (2008). Mobile agent migration modeling and design for target tracking in wireless sensor networks. *Ad Hoc Networks*, 6(1), pp. 1-16.

Zamir, A.R., Sax, A., Shen, W., Guibas, L.J., Malik, J., & Savarese, S. (2018). Taskonomy: Disentangling task transfer learning. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3712-3722.

Zhang, C., Xie, Y., Bai, H., Yu, B., Li, W., & Gao, Y. (2021). A survey on federated learning. *Knowledge-Based Systems*, 216, 106775.

Zhang, Q., Lu, J., Wu, D., & Zhang, G. (2018). A cross-domain recommender system with kernel-induced knowledge transfer for overlapping entities. *IEEE Transactions on Neural Networks and Learning Systems*, 30(7), pp. 1998-2012.

Zhang, W., & Mei, H. (2020). A constructive model for collective intelligence. *National Science Review*, 7(8), pp. 1273-1277.

Zheng, L., Yang, J., Cai, H., Zhou, M., Zhang, W., Wang, J., & Yu, Y. (2018). Magent: A many-agent reinforcement learning platform for artificial collective intelligence. In: *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 32, No. 1), pp. 8222-8223.

Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., ... & He, Q. (2020). A comprehensive survey on transfer learning. *Proceedings of the IEEE*, 109(1), pp. 43-76.

A1.9. References related to Chapter 8.

Alvargonzález, D. (2011). Multidisciplinarity, interdisciplinarity, transdisciplinarity, and the sciences. *International Studies in the Philosophy of Science*, 25(4), pp. 387-403.

Ashby, I., Caskurlu, S., & Exter, M. (2018). Evolving roles of faculty at an emerging hybrid competency-based transdisciplinary program. *The Journal of Competency-Based Education,* 3(1), e01059.

Balsiger, P.W. (2004). Supradisciplinary research practices: History, objectives and rationale. *Futures*, 36(4), pp. 407-421.

Bergmann, M., Jahn, T., Knobloch, T., Krohn, W., & Pohl, C. (2012). Methods for transdisciplinary research: A primer for practice. Campus Verlag, Frankfurt.

Boardman, C., & Ponomariov, B. (2014). Management knowledge and the organization of team science in university research centers. *The Journal of Technology Transfer*, 39(1), pp. 75-92.

Brown, V.A. (2015). Utopian thinking and the collective mind: Beyond transdisciplinarity. *Futures*, 65, 209-216.

Buanes, A., & Jentoft, S. (2009). Building bridges: Institutional perspectives on interdisciplinarity. Futures, 41(7), pp. 446-454.

Burgin, M., & Hofkirchner, W. (2017). Introduction: Omnipresence of information as the incentive for transdisciplinarity. In: *Information Studies and the Quest for Transdisciplinarity: Unity through Diversity*, pp. 1-7.

Cilliers, P., & Nicolescu, B. (2012). Complexity and transdisciplinarity - Discontinuity, levels of reality and the hidden third. *Futures*, 44(8), pp. 711-718.

CohenMiller, A., & Pate, E. (2019). A model for developing interdisciplinary research theoretical frameworks. *The Qualitative Report*, 24(6), pp. 1211-1226. https://doi.org/10.46743/2160-3715/2019.3558

Darbellay, F. (2015). Rethinking inter-and transdisciplinarity: Undisciplined knowledge and the emergence of a new thought style. *Futures*, 65, pp. 163-174.

Dawson, R. 2013. Building success in the future of work: T-shaped, Pi-shaped, and Comb-shaped skills. https://rossdawson.com/building-future-success-t-shaped-pi-shaped-and-comb-shaped-skills/, accessed on 11.01.2023.

Defila, R., & Di Giulio, A. (2015). Integrating knowledge: Challenges raised by the "Inventory of Synthesis". *Futures*, 65, pp. 123-135.

Fiore, S.M. (2008). Interdisciplinarity as teamwork: How the science of teams can inform team science. *Small Group Research*, 39(3), pp. 251-277.

Giri, A.K. (2002). The calling of a creative transdisciplinarity. *Futures*, 34(1), pp. 103-115.

Hadorn, G.H., Biber-Klemm, S., Grossenbacher-Mansuy, W., Hoffmann-Riem, H., Joye, D., Pohl, C., ... & Zemp, E. (2008). The emergence of transdisciplinarity as a form of research. In: *Handbook of Transdisciplinary Research*. Springer, Dordrecht, pp. 19-39.

Hoffmann, S., Pohl, C., & Hering, J.G. (2017). Exploring transdisciplinary integration within a large research program: Empirical lessons from four thematic synthesis processes. *Research Policy*, 46(3), pp. 678-692.

Horváth, I. (2004). A treatise on order in engineering design research. *Research in Engineering Design*, Vol. 15, pp. 155-181.

Horváth, I. (2015). An initial categorization of foundational research in complex technical systems. *Journal of Zhejiang University - Science A: (Applied Physics & Engineering)*, 16(9), pp. 681-705.

Horváth, I. (2017). A method for systematic elaboration of research phenomena in design research. In: *Proceedings of the 21st International Conference on Engineering Design*, August 21-25, 2017, University of British Columbia, Vancouver, Canada, pp. 1-10.

Horváth, I. (2022). Designing next-generation cyber-physical systems: Why is it an issue? *Journal of Integrated Design and Process Science*, (Preprint), pp. 1-33.

Horváth, I. (2023). A first inventory of investigational concerns for prognostic systems thinking based on an extended conceptual framework. *Journal of Integrated Design and Process Science*, (Preprint), pp. 1-25.

Horváth, I. Rusák, Z., & Li, Y. (2017). Order beyond chaos: Introducing the notion of generation to characterize the continuously evolving implementations of cyber-physical systems. In: *Proceedings of the International Design Engineering Technical Conferences & Computers and Information in Engineering Conference,* August 6-9, 2017, Cleveland, OH, ASME, pp. 1-14.

Horváth, I., & Pourtalebi, S. (2015). Fundamentals of a mereo-operandi theory to support transdisciplinary modeling and co-design of cyber-physical systems, In: *Proceedings of the International Design Engineering Technical Conferences & Computers and Information in Engineering Conference*, August 2-5, 2015, Boston, MA, ASME, pp. 1-12.

Jahn, T., & Keil, F. (2015). An actor-specific guideline for quality assurance in transdisciplinary research. *Futures*, 65, pp. 195-208.

Kläy, A., Zimmermann, A.B., & Schneider, F. (2015). Rethinking science for sustainable development: Reflexive interaction for a paradigm transformation. *Futures*, 65, pp. 72-85.

Klein, J.T. (2010). A taxonomy of interdisciplinarity. In: *Oxford Handbook of Interdisciplinarity*, Oxford University Press, Oxford, UK, pp. 15-30.

Lawrence, M.G., Williams, S., Nanz, P., & Renn, O. (2022). Characteristics, potentials, and challenges of transdisciplinary research. *One Earth*, 5(1), pp. 44-61.

Lawrence, R.J. (2015). Advances in transdisciplinarity: Epistemologies, methodologies and processes. *Futures*, 65, pp. 1-9.

Ledford, H. (2015). Team science. *Nature*, 525(7569), pp. 308-311.

Mâsse, L.C., Moser, R.P., Stokols, D., Taylor, B.K., Marcus, S.E., Morgan, G.D., ... & Trochim, W.M. (2008). Measuring collaboration and transdisciplinary integration in team science. *American Journal of Preventive Medicine*, 35(2), pp. S151-S160.

McComb, C., & Jablokow, K. (2022). A conceptual framework for multidisciplinary design research with example application to agent-based modeling. *Design Studies*, 78, pp. 1-13.

Melo, A.T., & Caves, L.S. (2020). Complex systems of knowledge integration: A pragmatic proposal for coordinating and enhancing inter/transdisciplinarity. In: *From Astrophysics to Unconventional Computation*, Springer, Cham, pp. 337-362.

Mobjörk, M. (2010). Consulting versus participatory transdisciplinarity: A refined classification of transdisciplinary research. *Futures*, 42(8), pp. 866-873.

NASEM (2019). Fostering the culture of convergence in research: Proceedings of a workshop. National Academies of Sciences, Engineering, and Medicine. National Academies Press, N.Y.

Nicolescu, B. (2014). Methodology of transdisciplinarity. *World Futures*, 70(3-4), pp. 186-199.

Nowotny, H., Scott, P., & Gibbons, M. (2003). “Mode 2” revisited: The new production of knowledge. *Minerva*, 41, pp. 179-194.

NRC (2015). Enhancing the effectiveness of team science. National Research Council. National Academies Press, New York.

Ochara, N.M. (2022). Reflecting on the nature of science and the 'little known' role of retroductive reasoning in research. Available at SSRN: https://ssrn.com/abstract=4154972. Accessed on 05. 01. 2023.

Osborne, P. (2015). Problematizing disciplinarity, transdisciplinary problematics. *Theory, Culture & Society*, 32(5-6), pp. 3-35.

Parker, J.R. (2008). Beyond disciplinarity: Humanities and supercomplexity. *London Review of Education*, 6(3), pp. 255-266.

Pohl, C. (2005). Transdisciplinary collaboration in environmental research. *Futures*, 37(10), pp. 1159-1178.

Rousseau, D. (2017). Systems research and the quest for scientific systems principles. *Systems*, 5(2), 25, pp. 1-16.

Schmidt, J.C. (2008). Towards a philosophy of interdisciplinarity. *Poiesis & Praxis*, 5(1), pp. 53-69.

Scholz, R.W., & Steiner, G. (2015a). The real type and ideal type of transdisciplinary processes: Part I - Theoretical foundations. Sustainability *Science*, 10(4), pp. 527-544.

Scholz, R.W., & Steiner, G. (2015b). The real type and ideal type of transdisciplinary processes: Part II - What constraints and obstacles do we meet in practice? *Sustainability Science*, 10(4), pp. 653-671.

Simon, D., & Schiemer, F. (2015). Crossing boundaries: complex systems, transdisciplinarity and applied impact agendas. *Current Opinion in Environmental Sustainability*, 12, pp. 6-11.

Stokols, D., Hall, K. L., Taylor, B. K., & Moser, R. P. (2008). The science of team science: overview of the field and introduction to the supplement. *American Journal of Preventive Medicine*, 35(2), pp. S77-S89.

Tanik, M.M., Gatchel, S., Horváth, I., Wan, T.H., Kim, K.-Y., Huang, J., Coatanea, E., Krämer, B., & Zeng, Y. (2021). Footsteps towards a transdisciplinary design and process science. *Journal of Integrated Design and Process Science*, 25(3-4), pp. 1-16.

Tebes, J.K., & Thai, N.D. (2018). Interdisciplinary team science and the public: Steps toward a participatory team science. *American Psychologist*, 73(4), pp. 2-34.

Tebes, J.K., Thai, N.D., & Matlin, S.L. (2014). Twenty-first century science as a relational process: From Eureka! to team science and a place for community psychology. *American Journal of Community Psychology*, 53(3-4), pp. 475-490.

Vogel, A.L., Hall, K.L., Fiore, S.M., Klein, J.T., Bennett, L.M., Gadlin, H., ... & Falk-Krzesinski, H.J. (2013). The team science toolkit: enhancing research collaboration through online knowledge sharing. *American Journal of Preventive Medicine*, 45(6), pp. 787-789.

Wiek, A. (2007). Challenges of transdisciplinary research as interactive knowledge generation - Experiences from transdisciplinary case study research. *GAIA-Ecological Perspectives for Science and Society*, 16(1), pp. 52-57.

Wiesmann, U., Biber-Klemm, S., Grossenbacher-Mansuy, W., Hadorn, G. H., Hoffmann-Riem, H., Joye, D., ... & Zemp, E. (2008). Enhancing transdisciplinary research: A synthesis in fifteen propositions. In: *Handbook of Transdisciplinary Research*, Springer, Dordrecht, pp. 433-441.

Yu, S., Bedru, H. D., Lee, I., & Xia, F. (2019). Science of scientific team science: A survey. *Computer Science Revi*ew, 31, pp. 72-83.

A1.10. References related to Chapter 9.

Bostrom, N. (2014). Superintelligence: Paths, dangers, strategies, Oxford University Press, Oxford, pp. 53-57.

Miller, W.D. (2022). The Future of Systems Engineering: Realizing the Systems Engineering Vision 2035. In: *Transdisciplinarity and the Future of Engineering*, pp. 739-747.

Ochara, N.M. (2023). Understanding digital economy dilemmas. In: *Proceedings of NEMISA Digital Skills Conference 2023: Scaling Data Skills for Multidisciplinary Impact*, Vol. 5, 2023, pp. 116–134.

Srinivasan, V. (2013). Why should we care about cyber-physical systems? *Journal of Computing and Information Science in Engineering*, 13(3), pp. 1-2.

Tang, Y., Liang, J., Hare, R., & Wang, F.Y. (2020). A personalized learning system for parallel intelligent education. *IEEE Transactions on Computational Social Systems*, 7(2), 352-361.

Appendix 2

Glossary

AGI	artificial general intelligence
AI	artificial intelligence
ANI	artificial narrow intelligence
ARM(s)	ampliative reasoning mechanism(s)
ASI	artificial super intelligence
ASRM	application-specific reasoning mechanism
AST	analytical systems thinking
BANGM	bits, atoms, neurons, genes, memes
CAM	conceptual advancement model
CES	cognitive engineering of systems
cisp*CPS(s)	complexificated, intellectualized, socialized, and personalized CPS(s)
CPAS(s)	cyber-physical agricultural system(s)
CPCS(s)	cyber-physical care system(s)
CPMS(s)	cyber-physical medical system(s)
CPmS(s)	cyber-physical military system(s)
CPNS(s)	cyber-physical nautical system(s)
CPPS(s)	cyber-physical production system(s)
CPA	cyber-physical augmentation
CPS(s)	cyber-physical system(s)
CPSoS(s)	cyber-physical system(s) of systems
CPSS(s)	cyber-physical-social system(s)
CPTS(s)	cyber-physical transportation system(s)
CRM(s)	computational reasoning mechanism(s)
DOS	designed operation space
EOS	external operation space
ESS	extended system space
i*CPS(s)	intellectualized cyber-physical system(s)
I4.0	industry 4.0
I-CPS(s)	intelligent cyber-physical system(s)
IEP(s)	intelligence exchange packet(s)

IES(s)	intellectualized engineered system(s)
IOS	internal operation space
IoT	Internet of things
ISS	initial system space
IT	information technology
KBS(s)	knowledge-based system(s)
MEMS(s)	micro-electro-mechanical systems
NASEM	National Academies of Sciences, Engineering, and Medicine
NEMS	nano-electro-mechanical systems
NG-CPS(s)	next-generation cyber-physical system(s)
NRC	national research council
OOS	optimal operation space
p-CPS(s)	personalized cyber-physical system(s)
PFP	paradigmatic feature profile
PSF(s)	paradigmatic system feature(s)
PST	prognostic systems thinking
RSS	reproduced system space
s*CPS(s)	smart cyber-physical system(s)
SAI	super artificial intelligence
s-CPS(s)	socialized cyber-physical system(s)
SSI	synthetic system intellect
SSK	synthetic system knowledge
xG-CPS(s)	x-generation cyber-physical system(s)