HOW TYPICAL IS YOUR PROJECT? THE NEED FOR A NO-MODEL APPROACH FOR INFORMATION MANAGEMENT IN AEC

SUBMITTED: June 2022
REVISED: November 2022
PUBLISHED: January 2023
EDITOR: Robert Amor
DOI: 10.36680/j.itcon.2023.002

Tamer E. El-Diraby, Professor & Director
Intelligent Buildings Digital Twinning Center, Dept. of Civil & Mineral Engineering, University of Toronto
Tamer.Diraby@utoronto.ca

SUMMARY: *This paper discusses the merit of using a no-model approach (no common product models or ontologies, etc.) for managing information in the AEC. It proposes an option for such an approach through the generation and analysis of semantic and social networks of communication between project stakeholders. The proposed approach advocates for a bottom-up discovery of knowledge constructs from stakeholder communication. Knowledge constructs are mini two-mode networks containing, on the one hand, clusters of concepts that appear frequently in the semantic networks of stakeholder communication; and, on the other hand, the social networks of stakeholders discussing these concepts.*
Using common models (such as IFC) has several limitations, including inflexibility to recognize and accommodate project contexts (which vary constantly), inability to timely capture the emergence of knowledge, and the scope creep problem (the ever-existing need to add more concepts to the common model from within and outside ACE domain). The no-model approach presented here is meant to complement and not replace the established model-based approach.
This approach is built on the belief in the ontological agency of project stakeholders: knowledge is a social phenomenon that emerges through interactions between people. It advocates a shift from a top-down format where experts or standards clearinghouses tell (force) practitioners what should be true about their project. In every project, stakeholders customize (the structure of) established knowledge and adopt elements from emerging knowledge to address project-specific needs. They use the more superior intelligence (the human one) to innovate a 'model of what they know' to guide the management of the project in a manner specific to its context. By studying projects' communication, we tell (inform) project stakeholders what knowledge constructs can be found in their communication. Unlike generic/static models, the resulting knowledge constructs are by default sensitive to project conditions.
We should re-design our information management systems to be able to recognize and adaptively use the constructs established by project teams to facilitate their sharing of data (along with the established scheme, such as IFC). Relatedly used constructs can be nominated as AEC-wide prototype constructs, representing what we know about a typical project. At the initiation of a new project, these can be the starting scheme used by information and communication systems. As the project evolves and the project's own constructs are generated, the project-specific constructs should guide the flow of information. Contrasting project constructs against prototypes should inform the stakeholders of not only what is factual about their view/model of knowledge, but also how unique are they (from generic/base knowledge).
This approach to no-model thinking is advantageous for several reasons. First, addressing the model rigidity problem. Because of the increasing complexity of projects, no single/standardized model can capture all contexts. Second, the increasing need for handling project unstructured data. The proposed approach helps formalize knowledge constructs from such data using network science. Third, recognizing and tracking the evolutionary nature of knowledge. Fourth, supporting innovation: instead of forcing knowers (people) to comply with a static model of reality, the new approach encourages them to imagine new possible futures/ worlds—after all, the true essence of digital twinning is to virtualize futures not just to digitize the present.

KEYWORDS: *no-model approach, knowledge constructs, IFC, information models, interoperability, semantic networks, social networks, evolutionary systems*

REFERENCE: *Tamer E. El-Diraby (2023). How typical is your project? The need for a no-model approach for information management in AEC. Journal of Information Technology in Construction (ITcon), Vol. 28, pg. 19-38, DOI: 10.36680/j.itcon.2023.002*

1 INTRODUCTION AND MAIN CLAIMS

Should we rely solely on models to capture knowledge, manage information and commonly represent data? On the one hand, this can make sense: after all, '*all projects are projects*'. They all share a common mosaic of concepts, hence the rationality/lure of common models. On the other hand, '*every project is unique*'. This is typically attributed to the difference in designs, stakeholders and their interrelationships, and resources available/used as well as, most importantly, the project context and boundary conditions. Within this evolutionary environment, (conformance to) static models can constrain (box) our ability to recognize the special needs of projects and to learn/capture new knowledge generated in each project. In short, while modeling is an intuitive concept, no model can be as complex as reality. So, "all models are wrong [and] the scientist cannot obtain a "correct" one by excessive elaboration. On the contrary, following William of Occam, [we] should seek an economical description of natural phenomena (Box 1976)".

Over the last four decades, Building Information Modeling (BIM), the leading modeling initiative in the AEC industry, has been predicated on the belief that we can agree on a conceptualization of our knowledge and use that to create and constantly upgrade a common information models and/or representation of (product) data. IFC (Industry Foundation Classes), the standard for product data exchange in the industry, is a structured, top-down, expert-driven model.

IFC context and initial scope were directed to the digitization of product (technical) data, at the building level and targeted mainly professionals. The needs of practitioners are no longer limited to the exchange of technical data. To manage today's projects, they need to model business data and link project and enterprise data schemes. Equally, they see data as a resource that should be used to generate insights. In this world, unstructured data is increasingly becoming important. It contains valuable (tacit) knowledge about project and company business logic that can be used to build business intelligence tools.

Instead of interoperable CAD, today's professionals are interested in full-scale digital twinning. A digital twin, as defined here, is not a digitization of the built environment. It is a virtualization of physical-cyber-social aspects of the future(s) of a facility. It is a tool for collaborative prescriptive (or even generative) modeling of how a building will be designed and run. Digital twinning is a venue for stakeholder engagement and the creation, exchange and application of diversified knowledge(s). Beyond information exchange and data interoperability, a digital twin encompasses managing unstructured data and developing business intelligence analytics—both requiring a deeper and more flexible representation of knowledge.

Should we expand our standardized data models, such as IFC, to be the (only) basis for digital twins? Expansion of IFC is not wise, given its increasing size and the scope creep problem: where to stop, what to include? More importantly, IFC has a restrictive 'ontological model' that is geared towards the technical aspects of (mainly) products—particularly modeling location, geometry and material data. It cannot be easily expanded to accommodate modeling unstructured data, let alone standardize its representation. Does developing ontologies solve the problem? The plethora of ontologies that was created in the domain over the last two decades indicates that they are not the solution too. In addition to the difficulty of agreeing on a common ontological conceptualization of the world, ontologies face three key challenges (see Herman 2019): context, scope, and knowledge evolution. No one ontology can accommodate the needs of knowledge capture and/or data modeling in all situations. It is hard to set the scope of an ontology in a way that can support modeling the expanding number of disciplines that are increasingly being used in AEC systems. Finally, the uncertainty of the future means that constant updates of ontologies are needed to capture the emergence of knowledge (Ding et al. 2006). This is still the case despite the promise of probabilistic relationships as means for accommodating evolution in ontologies (see, for example, Munch et al. 2022)

Maybe the solution is to re-consider the 'mentality' of modeling, including its associated commitment to top-down thinking; experts as knowers; conformance to standards; parametrization; and the dream of agreeing on a common ontological conceptualization that is valid across all contexts. To this end, this position paper aims to explore a no-model approach for capturing domain knowledge in the AEC industry, and subsequently using that to manage information and exchange data. The aim is not to replace IFC (and other structured information models). Rather, it is to complement them.

The arguments made here are rooted in phenomenology, which promotes freedom from static models. We need to observe and learn from actual practices instead of solely relying on expert-driven, top-down modeling (Turk 2001).

The arguments are further based on criticism of logical positivism's pursuit of 'universal slang' and foundationalism. Rejecting standardization is not intended to dismiss the formalization of the language we use to describe our knowledge. Rather, avoid the rigidity that can result from an over-belief in the existence of a complete model, and, worse, that we can reach such a model! The arguments are further made with an embrace of the value of constructivism in assigning ontological agency to practitioners and in supporting innovation.

Finally, the arguments and the proposed new approach are made with a deep commitment to Popper's falsification principles. The author has convections (what Karl Popper calls 2nd World) about the nature of knowledge (what he calls 1st world) that are presented here in an attempt to build established agreement about the problem and its solution (what he calls the 3rd world). As such, the proposed approach is predicated on inviting falsification.

It is important to note that the arguments made here are about models that claim to capture reality universally and are meant as a standardization mechanism for information exchange. Other types of models are not related to the topic of this paper, including, for example, those used for training, supporting decision-making software or capturing the requirements of a new project or the workflow of a process or the profile of a stakeholder. These models are valuable to the operations and management of information systems. In fact, there is value in working on models that aim to capture reality. They should not be dismissed altogether, as long as we recognize that they are falsifiable and are not to be used to the exclusion of no-model approaches.

To streamline the presentation of the ideas in this paper, background material was placed in an appendix—the interested reader can start by reviewing it. The Appendix summarizes some key concepts related to modeling, knowledge capture and epistemology.

This allows the next sections to jump directly into discussing the key claims of this article. The first part of the article discusses the motivation for the proposed approach, including summarizing the key drivers for considering a new approach for information management and knowledge capture in AEC. The argument centers around the suitability of standard models to the needs of AEC in the era of digital twinning.

This is followed by an overview of the proposed approach, which is presented early in the paper to help put later discussions in perspective. A summary of the theoretical foundations of the proposed approach is then presented in the subsequent section. The aim here is to put the model-based approach in context and contrast it to other possible approaches. The section provides arguments for the value and feasibility of a no-model solution that relies on using stakeholder communication as a source for capturing knowledge and on using semantic and social networks as means for formally representing and exchanging it. The second part of the paper provides details about the mechanics of the proposed approach, including discussions about how network analysis can be used to extract knowledge constructs and the use of these to suggest prototypical constructs that can be used to guide future projects. A second appendix is provided to summarize specific cases where elements of the proposed approach have been implemented.

2 PART I: MOTIVATION AND NEED FOR CHANGE

2.1 Point of departure

The most important paradigm of the proposed approach is that we can capture and manage knowledge (and information) without common models (or with limited needs for them). The 'common model' mentality faces the challenge of agreeing on shared conceptualization. It also faces the standardization-contextualization dilemma: there is a never-ending need for customizing the rigid structure of common models to match the always-changing specific conditions of each project (see Appendix I for more discussion). Even a hypothetical project with the same design and delivery schema repeated within similar contexts will have different knowledge constructs. This is, simply, because our knowledge evolves from project to project—we learn; and technology progresses. Contextualizing/modifying common models is a key task for practitioners (Hartman and Amor 2017). In fact, when searching for information, project stakeholders do not want to be limited to the confines of a model. They always want to receive and use contextualized information (Kraaijenbrick 2007). They want to 'know' and learn about the position/behavior of a building component in the project life cycle (Demian & Balatsoukas 2012). In other words, we need to recognize, capture and represent the specific context of projects when we manage AEC information.

The long history of classification systems, data modeling, and ontology development has identified (named) extensive lists of concepts. There is limited disagreement about this mosaic of concept names. The challenge in modeling is in describing their behaviour: the interrelationships between concepts; and the rules (axioms) that govern their definition and morphism. Why cannot concepts have variations in their definition, from which the user can select? For concept relationships, should they always be permanent/static. They can be probabilistic and contingent. In other words, our representation of knowledge (and models of data/ information) should avoid the tendency for generalization and reliance on rigid logic (foundationalism).

At the core of foundationalist thinking is the belief that we need formalized knowledge representation to support reasoning and we need to comply with data models to support (full) interoperability. But as of the 1990s, computers shifted from being reasoning machines into being communication amplifiers (Levitt 2007), hence the Internet. Growing out of (the limitations of) the reasoning mentality, our approaches to managing information and capturing knowledge should expand and change. First, recognize the importance of the communication-knowledge nexus. Communication between humans is not only a reflection but indeed an externalization of their knowledge. By adequately analyzing the interactions of project stakeholders, we can distill some of their knowledge. To this end, we must recognize the ontological agency of stakeholders as the source of contextualized knowledge, in contrast, to solely relying on expert modeling. As an epistemology, constructivism (more accurately social constructivism) maintains that knowledge is a creation of its community which observes and seeks to create mental constructs to explain sensory experiences. Consequently, knowledge is a relative construct situated within the socio-cultural context of a community of practice; and is bound to evolve over time and place (Gergen, 1995).

Second, data collection and information management should be integrated with/linked to communication management. Communication is a tenant of human interactions and a cornerstone in project management practices. Stakeholders are motivated to communicate extensively during a project. If the information management system is separated from daily communication, it will be a burden and may not be effective. As cases in point, take for example the low efficiency of developing as-builts, daily site logs, documenting design rationale and assumptions, and, indeed, the very task of populating and updating BIM models. If we device means to embed information management in project communication systems, stakeholders will, directly or indirectly, contribute to recording and updating project information effectively.

2.2 Overview of the proposed approach

Before digging into the rationale/details of the proposed approach, the summary below provides the reader with an overview—this can contextualize the remaining sections. As shown in Figure 1, the proposed approach advocates bottom-up discovery of knowledge through capturing and analyzing practitioners' communication (documents and chats). Using manual or automated means, transfer text corpus into semantic (concept) networks. The network is an objective way to formalize concepts and their relationships as they are discussed by project stakeholders (in contrast to free text, for example). Using network analysis measures, we can detect clusters of concepts: a collection of coherently connected nodes.

In addition to the traditional concept-to-concept investigation (in the semantic network), the analysis of concepts should also be investigated from the actor-to-actor view. In other words, we should examine the semantics of the concepts used at the individual, interpersonal and collective levels (Yang and Gonzalez-Bailon 2017). Using network measures, we can observe key concepts discussed by most stakeholders or by the most diversified stakeholders or the most influential actors.

To this end, a knowledge construct is defined in this paper as concept-to-concept-to-actor clusters. More accurately, a two-mode mini network of interconnected concepts linked to their interacting actors. The reality of (AEC) knowledge is complex, contextual, and evolutionary, which makes capturing it hard (Popper's 1st World). The most qualified knowers are the project stakeholders (particularly practitioners) who hold convections and learn from old experiences (Popper's 2nd World). As they manage the project they co-learn, negotiate and co-create situated (perspective on) knowledge constructs. When they communicate, they express this knowledge (Popper's 3rd World). Representing their exchanges in the form of social and semantic networks exposes the structure of their conceptualizations. The triangulation of social and semantic network analysis enables formalized extraction of their assertions/theorization: their own model of knowledge.

Of course, these knowledge constructs evolve over the project life cycle. As explained in Appendix I, co-creation is used here in the same sense currently used by sustainability transition and urban planning researchers, among others. In this view, the engagement of users is not limited to listening to them, sharing the solution development with them, or handing them the power to make decisions. Co-creation recognizes the knowledge of stakeholders as equally valid and valuable to that of experts. They lead the capture and representation of knowledge and the innovation of new solutions (see Frantzeskaki et al. 2018; Horsbøl 2018). In fact, the process of empowering stakeholders to innovate and generate new knowledge is seen by many as more important than the development of the solution itself (see Bucci and El-Diraby 2019).

As shown in Figure 1, concept clusters repeated across several projects can be nominated as prototypes or reusable clusters. A porotype construct represents a typical, yet adjustable, knowledge construct. They morph as knowledge evolves and as project context changes. In new projects, its knowledge constructs should be identified periodically. They should be compared to the prototypes. By studying the difference between the constructs of an ongoing project and the porotypes, we can advise project stakeholders about not only what is true about their project and what they agree upon, but also how different is that from typical projects. Equally important, observing the evolution of the constructs of a single project over its lifecycle presents a great opportunity to examine the innovative way stakeholders adapt typical constructs to the conditions of the project, change them, and/or blend them with new experiences.

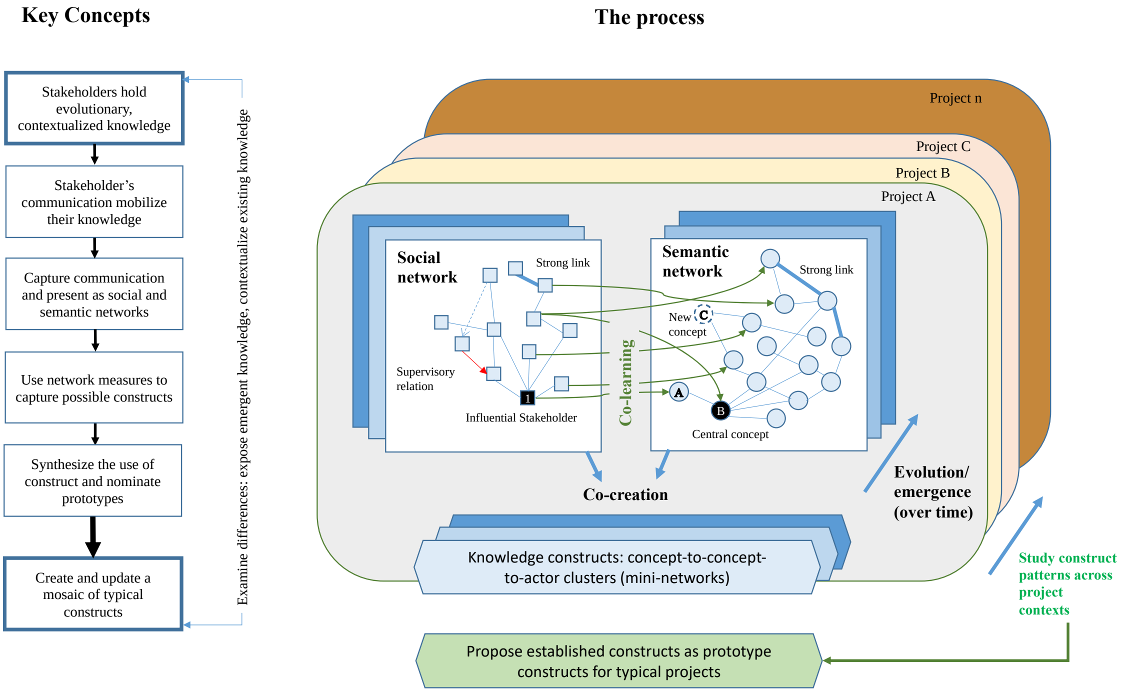

Figure 1: Generating knowledge prototype constructs through analyzing social and semantic networks

Instead of focusing only on conformance to a model of universal truths, we should embrace capturing and learning from uniqueness. That is, top-down models (such as IFC) force users to twist what they know to comply with standardized language (that does not necessarily reflect their views or project conditions). If we invest in capturing their own knowledge constructs, we can provide valuable service to stakeholders and, at the same time, learn from their knowledge, which is by default project-oriented and context-aware. The more we repeat this evolutionary 'learning-from-practice' process, the more we will be able to find meaningful candidate constructs that can serve as prototypes/benchmarks for future projects.

A major consequence of this approach is that we should re-develop information and communication modeling tools/systems to be able to recognize and re-use the knowledge constructs generated by each project. This, of course, does not mean that these tools should abandon relying on or using established standards. Standard models (such as IFC) are in fact a collection of prototypes. Integrating them in the management of information and communication systems brings technical knowledge constructs to the fold. As stakeholders customize these constructs, we could report to them the extent by which they have re-used or re-configured these standards.

2.3 Scope of theoretical arguments

Obviously, the circumstantial nature of the proposed approach could cast it as an unstable or inconsistent representation of knowledge. In the following section, I try to make the case that this type of representation is consistent with the nature of our knowledge, which is not only relativistic, evolutionary, and contextual but even, in many cases, contested. It is a common belief that among the main value propositions of practitioners in the AEC industry is their mastery of devising innovative work plans to meet the specific challenges of each project based on their collective knowledge. The competitiveness of contractors stems from their ability to re-use or embrace new knowledge to build adaptive systems for project execution as it evolves, especially as it deviates from plans, undergoes unforeseen changes or faces the ever-repeated unknowns. In fact, they are always motivated to plan things differently to avoid previous mistakes and/or learn from or use new tools. Within such a dynamic information environment, it is not appropriate to offer standardized models of only structured data as the basis for systems that are meant to serve them.

To this end, a few issues should be revisited (see Appendix I for more discussions):

- The role of experts, developers of information systems and researchers: they should complement their role as modelers with a new role as agents for knowledge discovery and brokers for its formalization. In other words, are they the arbitrators of what should be believed or are they the generators of falsifiable representations? Is the aim to lock knowledge or to promote co-learning and capturing the evolution of knowledge? Are we building classes that are meant to be stable? Or are we suggesting prototypes that can be used for benchmarking?
- Epistemology: evaluate the suitability of positivism (of standardized models) to the evolutionary and phenomenological nature of our knowledge. Question the (negative) implications of adhering to structuralism and foundationalism and contrast that to the possible benefits of coherentism. Putting users at the core of knowledge capture, epistemic contextualism emphasizes the knower context in evaluating knowledge claims. Epistemic relativism goes further by judging the truth values of knowledge claims based on the frames of reference considered when they were assessed (Egan et al. 2005). The difference in perspective and context of the knowers could be the reason why it is elusive to agree on a single common model of data or an ontology.
- Knowledge capture practices: fashion the process of knowledge capture to emphasize a balance between analysis and synthesis research methods. Using modal logic as a more flexible approach to capture rules than our traditional reliance on propositional logic.

If we imagine that there is a common mosaic of concepts (entities, features, attributes, or conditions) that underlays project conceptualization. There will be limited disagreements about the concepts in the mosaic or their names. The challenge is in understanding how the concepts are related and arranged (maybe this is why there is less disagreement on classification systems compared to ontologies). (See Wilde's 2020 discussion on relationships in ontology and his interesting views on the "death of objects"). More accurately, a new project is a new world that is created by tweaking and/or rearranging concepts. Several worlds can and do co-exist—because humans have competing views of reality. Beyond accepting this dualistic view, we should actually help practitioners theorize or imagine new worlds: futures or innovative states of the world that are not materialized yet. Further, if while negotiating or co-creating the work plans they stumble on them, we should help uncover/formalize these (by extracting their constructs).

Figure 2 is an attempt to visualize some of the theoretical discussions above.

The two fundamental dimensions in Figure 2 contrast, first (on the right and left sides of Figure 2), the mode of our work: are we in a mode of learning-from-practice or not (i.e. in a mode of codifying a standard); and second (on the top and bottom parts of Figure 2), our position on ontological agency: a common model by experts exists or not (do stakeholders have the agency to describe existence, or should we avoid that?) The latter dimension demarcates between positivism, empiricism and naïve rationalism (on the upper half of Figure 2) and constructivism and phenomenology (on the lower half of Figure 2). The vertical dimension demarcates between analysis, certainty and simplification and resisting dualism on the one hand and synthesis, emergence, co-creation and embracing dualism on the other hand. None of the lines or boxes in the chart are crisp. That is, there is a sort of fuzzy/gradient transition across both dimensions (This is why a color gradient is used for every quadrant and a common core is placed in the middle to denote that these concepts overlap).

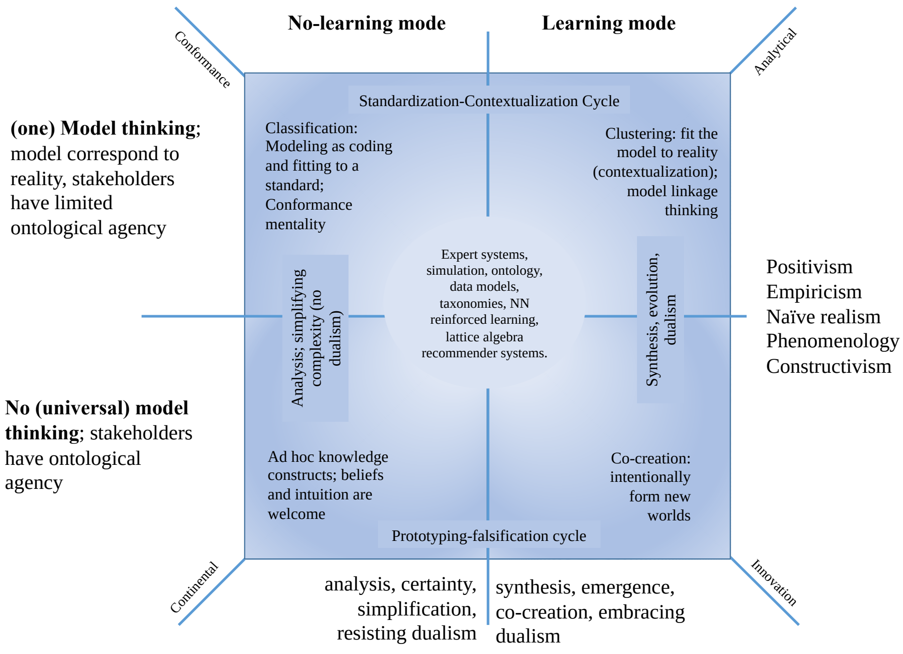

Figure 2: A view *of the knowledge capture world*

With that, there are four main quadrants:

Quadrant 1—the common model thinking and the mode of no learning: the archetype of this box is classification systems. The main purpose is to code/lock what we agree on into a structured format and to help fit new concepts into this structure.

Quadrant 2: the common model thinking and the mode of learning. This is a state where we want to upgrade or enrich a common (upper or abstract) model. This is where we use unsupervised learning (such as clustering and text mining systems) to find possible new additions or constructs (from a dataset or text corpus) to be addended or linked to an established structured model. To an extent, this quadrant accommodates a relativistic view of knowledge. In other words, we try to discover variations (or extensions) or contextualize some of our agreed-upon concept constructs, but resist changing the common upper model. This is similar to developing an MVD or the development of a new IFC update.

Quadrant 3: the no-common model thinking and the no learning mode: this is a brute constructivist view, where we aim to capture the knowledge of a group of stakeholders without aiming to generalize it. i.e., there is an aim for knowledge formalization, but no interest in re-using it. This is akin to a memoryless system that assumes the uniqueness of the knowledge of each project and resists the idea of knowledge reuse. It is an approach that resists creating or believing in a common ontology of reality.

Quadrant 4: the no-common model thinking and the learning mode: here we are closer to what can be called creative constructivism: stakeholders have an ontological agency, and we use it not just to create a model of their own world, but also to re-use elements of it. This is more of a memory-prototyping system. Knowledge is created at the project level. It can be re-used with modifications in future projects. Our role here is to support creativity. We accept dualism and emergence, and we rely on synthesis.

The proposed approach is in the 4th quadrant. A key feature in quadrant 4 is that there should be a focus on 'difference modeling'. In other words, in a new project, how did the stakeholders recall their previous knowledge and reformat it to fit the needs of the project or to express their views about how past experience can be used to serve the challenge of the world as it stands today? This stands in complete contrast to classification systems, which are all about conformance. This is depicted by a diagonal axis in Figure 2. The other diagonal axis differentiates between the traditional rejection of the personal experience (continental philosophy) and the analytical tradition. In representing or theorizing about knowledge, we are not limited to empiricist methods. We can accept, at least provisionally, the intuition of people as a source of new knowledge (to be falsified/examined by empiricist methods as Popper suggested), after all, project management is an art as well as a science. In other words, we can give experienced stakeholders the chance to try something new—an immensely needed attitude in an industry that has some of the lowest levels of R&D investments.

In the middle, the quadrants overlap. All quadrants can use or benefit from analysis/synthesis research tools such as expert systems, simulation, agent-based systems, and reinforced learning systems. A variety of means can be used to represent, and mobilize knowledge: taxonomies, ontologies, and IFC as well as semantic networks and recommender systems or even lattice algebra.

3 PART II: USING CONCEPT NETWORKS AND PROTOTYPING FOR A NO-MODEL APPROACH

The ideas proposed below are more about the "process" of using project communication to support capturing knowledge, and, in turn, support data and information management. In short, the main suggestion presented below is that the task of researchers, developers and service providers is to tell (in the sense of reveal, show, make aware, inform, notify, *narrate to*) project stakeholders what is factual about their project not to tell (in the sense of dictating, ordering, *prescribing to*) them how it should be. Our task is to capture and formalize, not enforce standardization of knowledge. Our challenge is to act as facilitators, not gatekeepers.

It is argued that models (like IFC) do not formalize reality. They reflect the consensus views of the modelers' community. "Therefore, several correct but different models may and should exist. Future software architectures in AEC should not be built on a unified, centralized model but on a combination of models, which may not be standardized but whose schemas are encoded in a standard manner (Turk 2020)". In constructivist thinking, the last part of this visionary view raises two questions: First, who creates these diverse models; are these *a priori* developed models? Second, what type of *standard encoding scheme* can accommodate this organic view of knowledge representation?

For the first question, the proposed approach targets the (Foucault's) vicious cycle of knowledge and power. Experts decide what is true and encode that into a model. Practitioners are to be confined to the use of this model. Furthermore, with time, the power of the model can overwhelm the expert modeler (e.g., new researchers) and tunnel their research and updates of the model to conform to the fundamental conceptualization used in building the original model. Instead of standardizing bodies holding all powers of modeling, the proposed approach encourages increased reliance on action research as the means to extract knowledge representation from the practices of users themselves. While an *a priori* standard model is not enforced, a set of porotypes are collected as benchmarks to help guide stakeholders as they externalize and formalize their tacit knowledge.

For the standard encoding scheme, the proposed approach suggests using networks to represent both semantics of project communications and the interrelationships between knowledge creators (the project stakeholders). Network representation of concepts and actors matches the connected nature of our concepts and our actors, without locking these interrelationships into the strict structure of a model or an ontology. Concepts (nodes) and links are discovered and arranged organically. This formalizes the capture of concepts in an unsupervised manner. The network format allows us to use graph theory to find (and constantly update) clusters of concepts (and their related actors) as representations of knowledge constructs. These constructs are not only context-aware but also project-wide. With this, we are no longer presenting a product data model (IFC) or even supporting a building-level model (BIM), we are moving towards recognizing and establishing a project information model (PIM). At the cross-project level (industry or sector level), we can nominate repeated constructs to be prototypes for future projects. These porotypes are benchmarks against which constructs of new projects can be contrasted. They can be modified by stakeholders to fit project conditions. More importantly, they can be combined as building blocks for imaging new worlds (when we are building a digital twin to study, for example, innovative schemes for concept relationships).

In summary, the key role of advanced knowledge management in AEC is to provide contextualized knowledge conceptualization to support the needs of project stakeholders and to learn new knowledge from their practices. Knowledge management (in AEC) is about discovery, brokerage, and formalization, as much as standardization. As we accept the paradigms of constructivist action research that knowledge constructs should be discovered from the interactions of stakeholders in each project, we should not dismiss the existence of cross-project constructs.

All projects share a mosaic of common concepts; and while their interrelationships (and the axioms governing them) are context-dependant, there exists a set of stable and reusable constructs across all projects. However, these constructs should be considered as prototypes that are used to promote learning from (not conforming to) past experience and support examining or creating non-typical constructs (see Majumdar and Sowa 2009).

3.1 Layout of potential implementation

The following section aims to offer ideas about a world where the proposed approach is used for knowledge management and discovery, recognizing the importance and dynamic nature of project contexts. The proposed approach is an iterative process that is built based on the following elements (the numbers in Figure 3 correspond to the numbers below):

1. (on the top, left corner) The key principle behind this approach is that knowledge is a social phenomenon and project stakeholders have ontological agency. Situated knowledge is a collection of constructs that evolve through the exchange of ideas, and experiences between project stakeholders, mostly in the form of unstructured data. As such, communication between them should be a main target for analysis. Not all of us (mainly, researchers and providers of information and communication management systems) should be just modelers. We can serve practitioners by also being knowledge brokers, where our role extends to the discovery of knowledge from the unfolding experiences of projects.
2. (on the lower, left corner) To support the objective study of stakeholders' communication, we should transfer their exchanges (reports, e-mails, or even chats) in the form of social and semantic networks: who is talking to who about what? See Appendix II for examples of how this can be done from several data sources and using different methods.
3. Knowledge constructs: We should use network analysis measures to study social and semantic networks. Clusters of concept-to-concept-to-actor are the basic blocks of our knowledge representation. We should orient the analysis of social and semantic networks to examine several facets of the knowledge constructs: temporal nature (their behavior over time), location (their association with specific zones), key themes (that address project issues such as safety or budget), stakeholders (parties interested in the key concepts in a cluster, such as designers, contractors, or local community), etc.
4. Memory prototyping: Of particular importance is to notice repeatedly used knowledge constructs. These are constructs that are observed and typically exist across projects in a fairly stable structure. These should be nominated to act as prototypes. Unlike object-oriented mentality, prototype-oriented thinking considers constructs as a guide (indeed, a sandbox) for project teams to use. They are to be amended based on the evolution of stakeholder communication.
5. (on the top right corner) Creating worlds: The prototype repository can include simple and complex prototypes; commonly used or context-specific prototypes; and real-world or hypothesized prototypes. Some standardized models (such as IFC) can be seen as a collection of prototypes. Combining different prototypes can create worlds: a set of connected concepts that represent scenarios for worlds that existed, commonly exist or can exist. For example, Case #6 and #8 in Appendix II showcase the use of Boolean Algebra to add the semantic networks of two projects. Generative algorithms can be used to create and examine a multitude of these worlds. This opportunity for imagining and testing future possibilities is the true objective and value of digital twinning: supporting innovating and virtualizing possibilities in the digital world before implementing them in the real world.
6. Re-orienting information management tools. We need to re-develop information and communication management tools to support the above environment. They should be oriented to promote communication, facilitate the development of social and semantic networks, and, then, use these networks to guide the tools themselves. At the start of a project, users can use typical constructs to act as a prototype knowledge representation for these tools. As the project evolves and as networks are created and analyzed, progressively these prototypes morph to reflect the context of the project and the specific knowledge generated by its stakeholders. It is important to build links between IFC and the evolving project constructs. This way, the system can benefit from the structured format of IFC when possible. The Green 2.0 is an example of this (see Case#5 in Appendix II)

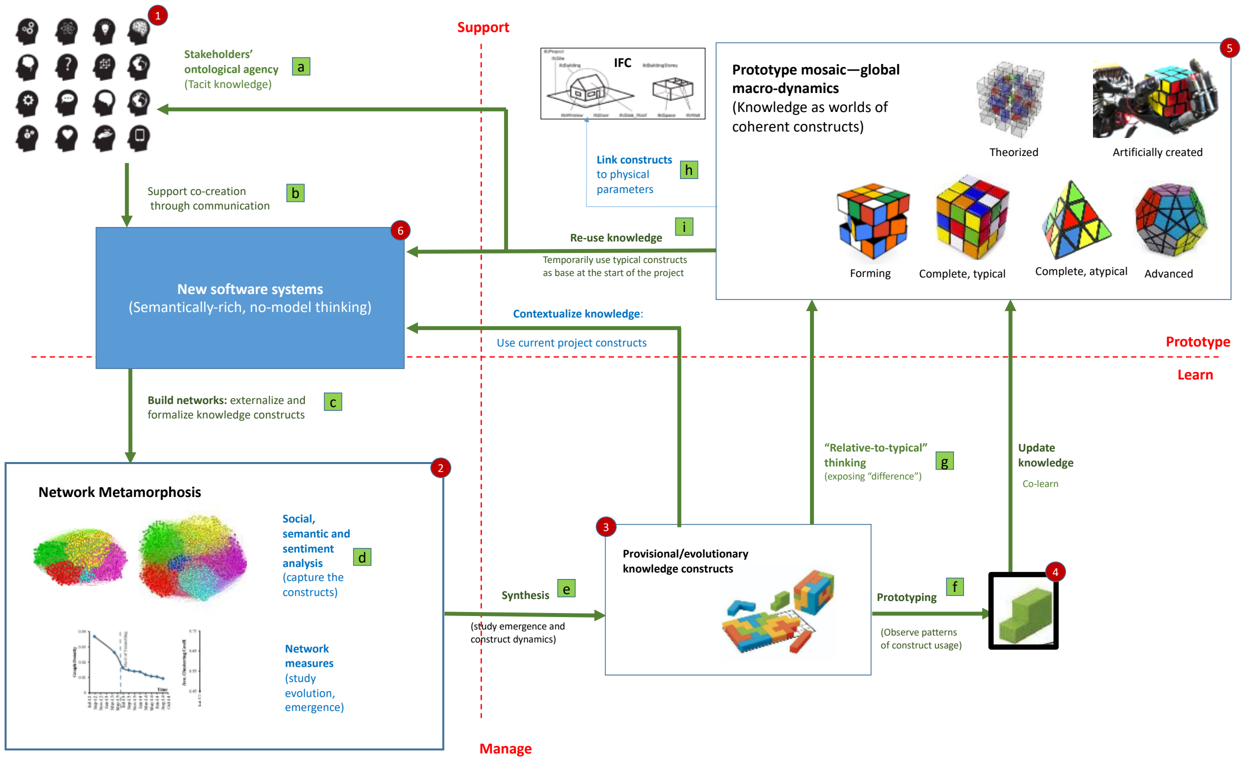

Figure 3: Outline of the proposed approach

The proposed system achieves the following features, listed on the arrows of Figure 3 and described below from the top-left corner in a contraclockwise direction (indicated by the green boxes):

a) Emphasize the ontological agency of stakeholders
b) Recognize that communication reflects co-created knowledge
c) Support externalizing stakeholder knowledge in the form of social and semantic networks
d) Use network science to study patterns in network structures
e) Deploy synthesis tools and analysis of microdynamics to observe the evolution of concept clusters
f) Observe cluster existence and re-use across projects and nominate repeatedly used clusters as a prototype
g) Instead of compliance to a common model, emphasize relative-to-a-prototype thinking
h) Exploit key features in established common models, such as using IFC to enrich prototypes and analysis and communication tools with geometry, location and material data.
i) Balance the need to re-use knowledge (by starting the project based on prototypes) with the need to contextualize knowledge representation (by updating base prototypes based on project networks).

3.2 The paradigm shift

The proposed approach is, essentially, a collaborative and continuous process for learning and for the discovery of knowledge constructs. The proposed approach is rooted in constructivism. However, it recognizes the value of positivistic thinking—mosaic prototypes should be linked to corresponding concepts in established models (such as IFC). While IFC is an exchange standard, linking it to prototypes means that knowledge constructs captured from user interaction can be fed into or used with OpenBIM platforms. This achieves a key goal of this proposal: making the link between information and communication systems seamless.

One of the most fundamental shifts that are advocated in the new approach is to not just reconsider structuralism, but also reconsider complete adherence to foundationalism. Traditional foundationalist views rely on a basic set of beliefs/rules from which other rules can be generated (using propositional logic). In the case of IFC, these basic beliefs are encoded in the common model through deterministic relationships between mosaic concepts. Any knowledge construct (emerging from a project) has to 'adhere' and be consistent with the established rules. This is a typical outcome of believing that a "*project is a project*" (they share a common/standard conceptualization).

In contrast, coherentism does not accept the existence of basic truths and establishes that rules should corporate each other. It holds that beliefs mutually support each other when they belong to the same coherent belief set. The epistemic justification for believing a proposition is a matter of fitting in or cohering with other propositions one believes. Instead of the pyramid-like structure of foundationalism, coherentism establishes a raft structure, where each new rule (or knowledge construct) is a plank that has to link to other parts of the raft. Interestingly, while coherentism requires local consistency between rules/truths, it can result in disagreement between non-localized truths—ones that are further away from each other. This type of coherentism, what Roche (2013) calls non-global coherentism, does not require universal consistency between truths. It accommodates the common view that "*each project is unique*".

A third approach, adopted mostly by this paper, foundherentism (Haak 1993), strikes a pragmatic mix. It rejects the uni-directional derivation of beliefs from a class of basic beliefs. Yet it allows for experiential evidence as a basis for other beliefs. It is possible that a set of *Actors* is more or less justified, at time *t* (during a specific project), in believing that *p (is truthful)*, depending on the quality of the instantaneous evidence supporting that. Metaphorically, knowledge constructs (or beliefs) are discovered much like a crossword puzzle: the clues of the puzzle are the experiential evidence, while the already existing words that intersect the missing word are the reasons for the belief. In the proposed approach, the first is the concept clusters extracted from the semantic networks of an ongoing project. The latter is the prototype constructs.

Pragmatically, this line of thinking fashions project information management after commonly used enterprise/business management practices. Organizations retain a set of key beliefs that distill common industry wisdom and/or encapsulate the identity of the organization or its business model. Complete alignment of various departments' needs/rules/beliefs (beyond the basic rules) is never attained. Rather organizations tend to focus on synchronizing a level of coherence across departments (Choo 2016).

In effect, this is a call for accepting a relaxed definition of concepts and relaxed or contingent relationships between them— a type of probabilistic justification or Bayesian Epistemology (see Morganti 2018; Huemer 1997; Olsson 2018). Modal logic can be a key tool in supporting the description of what is (should be) true about these constructs (Salama and El-Gohary, 2013).

A key concern about the proposed approach is the accuracy and stability of inferences made based on the constructs. No clearinghouse has checked and ascertained the logicality of constructs. And constructs change over the lifecycle of a project, which means that using them for inferences can yield different results. Perfect inferences and complete verification were reasonable targets in simpler times. For example, early on in the life of the Internet, Yahoo hired staff to manually create a web directory, where almost every webpage was annotated. Obviously, we now accept the "search approach" of Google with all its imperfect inference issues (Curry 2021). We need to change the traditional belief that we are managing a data platform (a database or a data warehouse), for which we need a data model into the recognition that, in today's environment, we are dealing with dataspaces (Franklin et al. 2005). The latter encompasses the generation, movement and exchange of data among actors with complex information supply chains along different levels of organizational hierarchies and within varying levels of sophistication of the work and information management environments (Curry 2020; Fernandez et al. 2020).

Two pragmatic questions are important here. First, are the systems/services and processes described above achievable in a meaningful way? Given the progress in other domains and recent work in the AEC, the answer is: most probably. Appendix II lists some work that has been done along the same lines as the proposed approach. Second, while the triangulation between the three types of analyses (semantic, sentiment and social) seems to offer valuable insights, it may raise privacy concerns. Several anonymization approaches can be used to help address this issue. However, the link between privacy protection and the value of analyzing communication patterns is an emerging industry-wide issue that we are facing and have to find solutions for. It is not uniquely created or caused by the proposed approach.

3.3 Semantic networks

The concept network is a representation of project-level knowledge. Project-level (in contrast to product-level) data/knowledge models can address the complex and context-sensitive nature of project information. For example, Hartman et al. (2009) proposed a project-centric methodology that relies on the ethnographic observation of practitioners. The methodology is aimed at tracking the requirements of project stakeholders and their iterative attempts to organize project information. (Note: in this proposed approach, ethnographic analyses are done through triangulation between social and semantic network analyses).

Relying on project-specific networks is a form of instance-based learning (sometimes called memory-based learning), which is a family of learning algorithms that, instead of performing explicit generalization, compare new problem instances with instances seen during training, which have been stored in memory (Daelemans and van den Bosch 2005). Examples of instance-based learning algorithms are the k-nearest neighbor algorithm and kernel machines. It is called instance-based because it captures constructs directly from the training instances themselves. One advantage that instance-based learning has over other methods of machine learning is its ability to adapt its model to previously unseen data. Instance-based learners may simply store a new instance or throw an old instance away.

In addition to being project-specific, semantic networks have five main advantages (Appendix II includes cases showcasing these advantages). First, just the fact that semantic networks provide a visualization of project concepts is a major advantage. In analyzing patterns of information search by engineers (on the web), Kraaijenbrink (2007) found that they prefer to browse tree-based hierarchical information systems both downwards and upwards. This enables them to understand and explore the context of a specific information item.

The second main advantage of semantic concept networks is the value/opportunity of the application of graph theory to analyze concept interconnections. A semantic network is very close to an ontology: it includes a set of concepts connected through relationships. Using network analysis measures, we can study the behavior and the meaning of the concepts. In addition to centrality, clustering techniques can be used to discover clusters. Network-level measures help observe the evolution of concept relationships within the same project. If we track these measures at several stages of their life cycle, then you can see the dynamics of the context and the evolution of knowledge. One key feature is that the network automatically includes a representation of boundary condition concepts (see Case #4 in Appendix II).

The third main advantage of using a concept network as a representation of knowledge constructs is the ability to objectively conduct cross-project comparisons, by using network similarity measures. Contrasting a project network to another can show the differences between projects. Contrasting it to historical projects can capture some of the evolution of knowledge. Contrasting a project network to a baseline prototypes can help measure how unique a project is. i.e., an organization can develop a baseline network (with its own constructs) to act as a benchmark. Deviations from this benchmark are an indicator of uniqueness. The baseline network can be built in a variety of ways (see Appendix II). It could be developed through text analysis of a large number of documents. For example, term frequency methods (or text retrieval methods in general) can be used to detect the unique terms and then their association. It is possible to build the base network by surveying a set of experts (top-down approach). It is also possible, and better, to build it through a Boolean addition of all (or some) project networks. This is another major feature of expressing knowledge as a network: you can generate variations or extensions for them by adding them (see Case #6 and Case #8 in Appendix II).

Fourth, tracking patterns in networks can detect clusters that can serve as prototype knowledge constructs. One of the key similarity analysis approaches in networks is blockmodeling. The typical similarity measures do so through matching nodes: if two networks have several similar nodes then they are similar. Nodes are judged as similar if they are connected to the same other nodes in both networks. Instead of focusing on the node level, blockmodeling aims to find blocks of nodes that are repeated in two networks. It iterates finding a set of coherent mini networks repeatedly used within several large networks. If these small networks (blocks) are reduced into nodes, then the main (typically, complex) network can be simplified. In other words, instead of studying multitude of sporadic nodes, we examine blocks of them. Blockmodeling is a tool for learning. Blocks that are repeated across many projects can be good candidate knowledge constructs (prototypes) to be added to the organizational baseline network or to the AEC concept mosaic (see Case #7 in Appendix II).

Fifth, the capture of unstructured data into a network format allows for linking the network concepts to IFC objects (see Case #5 in Appendix II). This can be very beneficial to enriching IFC-based systems with concepts from unstructured data. Such linkage can help cross reference drawings/IFC files to project documents (within common data environments-CDE). In fact, the use of networks to capture a model of project data can also be applied to structured product data—specifically, IFC. For example, Khalili et al. (2013) presented a methodology to use topological models to represent spatial relationships among objects in a building based on their IFC representation.

Graph models have been used to simplify complex geometric computations (Lee and Kwan 2005). 3D objects are arranged in a network based on two main features: adjacency and connectivity. Each building element in the IFC file is represented as a node in the graph. Network-based analysis such as graph traversal algorithms can then be used to check, for example, constructability rules or code compliance. Linked IFC data provide any additional physical features of the products, such as placement, shape, length, and weight (see Nguyen et al. 2005; Van Treeck and Rank 2007).

Experience-based reasoning & prototyping:

Analyzing precedence is an established paradigm not only in artificial intelligence but also in cognitive science (Leake 2001). Case-based reasoning (CBR) has been used extensively in construction project management—after all construction is an experience-based industry (see Fong & Choi 2009; Dogra 2020). To enhance the formalized capture and retrieval of cases, recently, researchers combined CBR with several tools, including ontologies (Yurchyshyna and Zarli 2009;) and natural language processing (Wang et al. 2022). While the proposed approach is precedence-based, it differs from CBR in several ways. In the proposed approach, a project is a case that can be benchmarked by future projects. Unlike traditional CBR, where the case is captured in the form of text, project cases are captured as socio-semantic networks. Instead of using simple natural language processing to retrieve cases in traditional CBR, project cases are re-used based on network analysis. Instead of traditional text retrieval methods, viewing concepts as nodes in a network provide better chances for understanding their context and meaning (Firth's famous point about knowing words by the company they keep). Adding a set of rules or an ontology to a CBR system can enhance the outcomes of text retrieval methods. However, the evolutionary nature of cases means that the ontology (with its static structure) will not be able to recognize evolving knowledge captured in new cases—it will still reason about the new case based on the built-in rules. Ontology has to be scoped to the minimum conceptualization to make sure that we give the contents of cases more prominence in the retrieval process, which reduces the value of ontology (Hamdan and Scherer 2022). Retrieving cases based on contrasting network measures, in contrast to relying on static rules in an ontology, can support dynamic capture of concept evolution— with every project, the stored clusters morph. New clusters are matched to these updated ones.

Three additional differences between traditional CBR and the proposed approach are fundamental. First, the analysis of cases (more accurately the networks of each case) is not part of a decision support system, as is typical in CBR. Rather, an exercise in knowledge capture (more accurately, seeking prototypes). We encapsulate every new 'case' not just as a text corpus but as a network; what we capture is not just 'lessons learned' but 'knowledge constructs discovered'. Second, while the knowledge constructs can be used for reasoning, the main purpose of their extraction and use is to support building communication and information management tools that are more semantically rich and can recognize the context of an ongoing project (Step #6 in the proposed approach). Third, and most fundamentally, CBR is built on finding similarities. The proposed approach is built on finding differences: how different are the concept clusters of an ongoing project compared to established prototypes? To an extent, looking for similarity assumes that history repeats itself. Looking for difference engrains the belief in the incremental evolution of knowledge; highlights the importance of seeking/analyzing new knowledge.

3.4 Social networks

Knowledge (development) is a social phenomenon. To understand a message, the audience and the speaker must "share mutual concern and a common background. The unspoken is at least as important as the spoken part of a message (Miettinen and Paavola 2014)." The knowledge represented in the semantic networks cannot be separated from the knowers who developed it—especially as they continually revise it.

Typically, social network analysis in AEC focused on understanding person-to-person relations, communication flows, and representing project organization. In the proposed approach, we need to analyze social networks because it is not possible to parse knowledge (in the concept networks) without understanding the knowers (knowledge producers). How does the structure of the actor-network impact the development, structure and evolution of knowledge (concept networks)? This type of activity theoretical analysis has not been widely considered in BIM and AEC data modeling practices. Activity theory rejects behaviorism and embraces actor agency and the unpredictable nature of technological and social advancement in the evolution of knowledge. The motives, commitments, innovation and experimentation by actors are key elements to the production of knowledge (Hartmann et. al 2009).

Action research processes include an iterative cycle of observation, identification of problems, development of technical solutions, and implementation of the developed solutions (Eden and Ackermann 2018). This mode of phenomenological thinking has dominated field practices over melaena—the information systems supporting them may as well mimic that (Miles, et al 2002). However, action research methodology, in general, does not offer detailed tools and techniques to achieve such an understanding. This is why Hartmann et. al. (2009) proposed complementing action research with ethnographic research, which, in the era of co-creation and unstructured data, is gaining momentum among technology researchers. Ethnographic research requires frequent reviews and discussions of findings with project members. In the proposed approach, the evolution of topics and the formation of sub-communities can be one means to observe that (Rashid et. al. 2019).

Profiling the users is as essential as profiling concepts. Construction knowledge is often tacit. It is the result of individual experiences as much as it is the result of analytical theorization (see the diagonal dimension in Figure 3). So, we need to understand the actor networks to better interpret the results of semantic analysis. Social network analysis provides several advantages. First, the analysis of social networks can be viewed as means to examine the provenance of knowledge (see Case #2 in Appendix II). How was it developed, what are the qualifications of the knowers who developed it, and in what context? Users of information systems typically prefer to view general details about the entire retrieved document (i.e., metadata elements) after that they became interested in exploring document contents (Pharo 2008). This may refer to a mentality of "evaluate the source and investigate the contents". In the proposed approach, combining social network analysis with semantic analysis can help achieve that.

Second, the interpretation of network analysis is subjective, circumstantial and ever-changing. One of the most reliable interpretations of these analysis is that of project stakeholders. So, it is important to profile them. In an evolutionary system, this should not be seen as invaluable or confusing. Rather, it should be viewed in the context of creating possibilities: Some of the results can spot innovative ideas; some may capture an emerging practice.

Third, profiling actors is also important to the efficiency of any information system. With the increasing frequency of information exchanges and their complex multidisciplinary nature, new information systems have to embrace a mentality of service: delivering the right information to the right person at the right time. It is important therefore to read the minds of project stakeholders (see Castaño et al. 2017). This recommender-like style stands in contrast to standardized models, which have limited or no appreciation of actors. In fact, the lack of adequate profiling of actors' needs, can be a factor in the increasing dissatisfaction with IFC-based information systems among AEC practitioners (van Berlo, Natrop 2015). The role of technology has to shift to supporting project teams instead of telling them what the model dictates or, more troubling, forcing them to adjust their needs to match the model constructs (see Case#3 in Appendix II).

Some important caveats should be considered here. With project staff knowing that their communication contributes to creating knowledge constructs, they may tend to intentionally influence outcomes or even manipulate them. According to Grounded Theory (Glaser and Strauss 2017), we should always take into account this social phenomenon. In other words, in using a no-model approach to reduce the power of experts (who develop the standardized models), do we replace that with the power of influential project stakeholders? To address that, we can use two steps: 1) benchmarking project-specific semantic networks against the established relevant prototypes; 2) studying the social network to make the actor power balance transparent.

Sentiment analysis

Sentiment analysis can be a great addition to the proposed approach. Through socio-semantic analysis, we know the topics/concepts from the semantic network. We know the people who have generated them, are advocating for them, or are challenging them. Adding analysis of sentiment about concepts and associated with groups of stakeholders can help us move from finding topics to modeling opinion. In a simplified definition, an opinion is a view about a topic held by a person.

In other words, if we define an idea as a block of concepts, then we can observe how such an idea evolves. Does it spread like a rumour—no changes to the contents (everybody is passing the buck)? If the content of the block change, then people are contributing to it. Observing the changes in sentiment as an idea grows in depth and spreads across the network can help us detect how it evolved: through negotiation, co-learning, or the decision of influential actors (Zolin et al. 2004).

4 DISCUSSION: ANALYSIS-SYNTHESIS & STANDARDIZING-PROTOTYPING

At the product level, IFC has reached its limit in modeling structured data and is not suitable for managing unstructured data. The latter is assuming an increasingly important role in the era of digital twinning as it can support business intelligence analytics. Furthermore, because digital twinning is about virtualizing whole futures, our knowledge capture and information management systems must be scaled from the product level to the project level and even the asset life cycle level. In fact, an essential need for practitioners is to go a step further: linking project or asset information to the overall enterprise information architecture.

In this context, the big challenges in product-based modeling of AEC concepts are the uniqueness of projects in terms of the dynamics of their context; the difficulty of capturing subjectivity (especially of boundary conditions); and (the neglect of) the evolutionary nature of knowledge. The dominance of logical positivism, structuralism and normative thinking in product-based standards (such as IFC) has made them less effective in addressing these issues.

The approached presented in this paper advocates mixing the prevailing mentality of top-down modeling of concepts (of IFC and ontologies) with bottom-up discovery of knowledge constructs. These are concept-to-concept-to-actor clusters that are extracted from stakeholders' interactions. The proposed approach relies on collecting text from project sources (documents, e-mails, and chats), transferring their main concepts into semantic networks and, at the same time, capturing the social network of stakeholders' communication. Using socio-semantic network analysis (and sentiment analysis), we can extract possible falsifiable knowledge constructs. These constructs are akin to mini ontologies organically grown by project stakeholders. However, they are not limited to product modeling and are, by default, contextualized to project conditions. Our information systems must be re-developed to capture, recognized and adaptively use these constructs as foundational conceptualization of knowledge.

Through observing patterns and repeated use of these constructs throughout the project life cycle and across projects, commonly used ones can be nominated as prototypical constructs. When the socio-semantic networks of a new project are analyzed, its resulting concept clusters are to be compared to these prototypes. By observing similarities and differences to existing prototypes, we can evaluate the uniqueness of project constructs. We can examine what new or different conceptualization did stakeholders (the knowers) generate/ use.

Concepts are triangulated based on their location in the network, which can give better chances to understand their meanings. As the network evolves, the position and interrelationships of concepts reflect the changes in the project knowledge. Iterative and transparent analysis of the network over the project life cycle invites stakeholders to falsify the generated networks. Because of the formality of the networks, stakeholders can provide valuable criticism or corrections more objectively.

In addition to modeling, researchers and developers of information management systems have an equally important role: knowledge discovery and brokerage. Instead of just telling project stakeholders how they should conform their knowledge to a standard, we should also help them discover and negotiate the very structure of their knowledge. Our role is to discover the local micro-dynamics of knowledge constructs (within the specific context of a project) and scale that to the macro-dynamics: a mosaic of possible prototype constructs that can be used for benchmarking and discovering knowledge evolution/ emergence. "The importance here is the formation of stable global order, which is not fixed, periodical, nor chaotic, but complex structure. A stable global order gives a new function, and if the function meets the specified purpose, one can adopt it as a solution. The definition above is compatible with that used in Artificial Life and Complex Adaptive System studies. (Ueda 2001)".

Rejecting a common model should not lead to a sort of anilinism. In the new approach, common constructs still exist. But instead of being standards (for conformance), they are only porotypes to learn from. Capturing project networks formalizes conceptualizations, yet it is not a structured or top-down formalization. Rather, an evolutionary map of the results of deliberations by multiple project stakeholders. Instead of traditional expert-based knowledge sourcing, the source of knowledge in this proposal is the stakeholders.

REFERENCES

Björk, B. C. (1992). A unified approach for modelling construction information. Building and Environment, 27(2), 173-194.

Bonduel, M., Vergauwen, M., Klein, R., Rasmussen, M. H., & Pauwels, P. (2018). A novel workflow to combine BIM and Linked Data for existing buildings. ework and E-business in Architecture, Engineering and Construction, 12th European Conf. on Product and Process Modeling (ECPPM), Copenhagen, Denmark, 347-354.

Box, G. E. (1976). Science and statistics. J. American Statistical Association, 71(356), 791-799.

Bucci, M., and El-Diraby, T. E. (2018). The functions of knowledge management processes in urban impact assessment: the case of Ontario. Impact Assessment and Project Appraisal, 36(3), 265-280.

Capilla, R., Bosch, J., & Kang, K. C. (2013). Systems and software variability management. Concepts, Tools and Experiences, 10, 2517766.

Castaño, J. M., van Amstel, F., Hartmann, T., & Dewulf, G. (2017). Making dilemmas explicit through the use of a cognitive mapping collaboration tool. Futures, 87, 37-49.

Castaño, J. M., Hartmann, T., Dewulf, G. P., & van Huffelen-de Kort, I. A. (2015). What is going on and what should we do? Divergent frames in multifunctional projects. Engineering project organization, 5(1), 36-48.

Cemesova, A., Hopfe, C. J., & McLeod, R. S. (2015). PassivBIM: Enhancing interoperability between BIM and low energy design software. Automation in Construction, 57, 17-32.

Cerovsek, T. (2011). A review and outlook for a 'Building Information Model'(BIM): A multi-standpoint framework for technological development. Advanced Engineering Informatics, 25(2), 224–244.

Choudhary, A. K., Oluikpe, P. I., Harding, J. A., & Carrillo, P. M. (2009). The needs and benefits of Text Mining applications on Post-Project Reviews. Computers in Industry, 60(9), 728-740.

Checkland, P. (1995). Model validation in soft systems practice. Systems Research, 12(1), 47-54.

Cherneff, J., Logcher, R., & Sriram, D. (1991). Integrating CAD with construction-schedule generation. Computing in Civil Engineering, 5(1), 64-84.

Chinowsky, P. S., Diekmann, J., & O'Brien, J. (2010). Project organizations as social networks. Construction Engineering and Management, 136(4), 452-458.

Choo, C. W. (2016), The Inquiring Organization: How Organizations Acquire Knowledge and Seek Information, New York, NY: Oxford University Press,

Costa, G., & Madrazo, L. (2015). Connecting building component catalogues with BIM models using semantic technologies: an application for precast concrete components. Automation in Construction, 57, 239-248.

Curry E. (2020) Real-time linked dataspaces: Enabling data ecosystems for intelligent systems. Springer Nature.

Curry, E. (2021) From Data Platforms to Dataspaces: Enabling Data Ecosystems for Intelligent Systems. Keynote presentation. 9th Linked Data in Architecture and Construction (LDAC2021), Luxembourg.

Daelemans, W., & Van den Bosch, A. (2005). Memory-based language processing. Cambridge University Press.

Demian, P., & Balatsoukas, P. (2012). Information retrieval from civil engineering repositories: importance of context and granularity. Computing in Civil Engineering, 26(6), 727-740.

Ding, L. Y., Zhong, B. T., Wu, S., & Luo, H. B. (2016). Construction risk knowledge management in BIM using ontology and semantic web technology. Safety Science, 87, 202-213.

Ding, Z., Peng, Y., & Pan, R. (2006). BayesOWL: Uncertainty modeling in semantic web ontologies. In Soft computing in ontologies and semantic web (pp. 3-29). Springer.

Dogra, P. (2020). Reliance of the field supervisors on experience-based tacit knowledge and barriers to knowledge sharing. Doctoral dissertation, Purdue University.

Eastman, C. M., Jeong, Y. S., Sacks, R., & Kaner, I. (2010). Exchange model and exchange object concepts for implementation of national BIM standards. Computing in Civil Engineering, 24(1), 25-34.

Eden, C., & Ackermann, F. (2018). Theory into practice, practice to theory: Action research in method development. European Journal of Operational Research, 271(3), 1145-1155.

El-Gohary, N. M., & El-Diraby, T. E. (2011). Merging architectural, engineering, and construction ontologies. Computing in Civil Engineering, 25(2), 109-128.

Fernandez, R. C., Subramaniam, P., & Franklin, M. J. (2020). Data market platforms: Trading data assets to solve data problems. arXiv preprint arXiv:2002.01047.

Fong, P. S., & Choi, S. K. (2009). The processes of knowledge management in professional services firms in the construction industry: a critical assessment of both theory and practice. Knowledge Management, 13(2), 110-126.

Franklin, M., Halevy, A., & Maier, D. (2005). From databases to dataspaces: a new abstraction for information management. ACM Sigmod Record, 34(4), 27-33.

Frantzeskaki, N., Hölscher, K., Bach, M., & Avelino, F. (2018). Co-creating sustainable urban futures. A primer on applying transition management in cities, Future City, 11.

Gao, G., Liu, Y.-S., Wang, M., Gu, M., and Yong, J.-H. (2015). A query expansion method for retrieving online BIM resources based on industry foundation classes. Automation in Construction, 56, 14-25

Gergen, K. J. (1995). Relational theory and the discourses of power. Management and organization: Relational alternatives to individualism, 29-50.

Glaser, B. G., & Strauss, A. L. (2017). Discovery of grounded theory: Strategies for qualitative research. Routledge.

Gogineni, S., Lindow, K., Nickel, J., & Stark, R. (2020). Applying contextualization for data-driven transformation in manufacturing. IFIP International Conference on Advances in Production Management Systems (pp. 154-161). Springer, Cham.

Hamdan, A. and Scherer, R. (2022). An ontology-supported case-based reasoning approach for damage assessment. European Conf. on Product and Process Modeling (ECPPM), Trondheim, Norway.

Harman G (2019) Editorial Introduction for the Topical Issue "Object-Oriented Ontology and its Critics". Open Philosophy, 592–598

Hartmann, T., Fischer, M., & Haymaker, J. (2009). Implementing information systems with project teams using ethnographic–action research. Advanced Engineering Informatics, 23(1), 57-67.

Hartmann, T., Amor, R., & East, E. W. (2017). Information model purposes in building and facility design. Computing in Civil Engineering, 31(6), 04017054.

Hietanen, J., Lehtinen, S. (2006). The useful minimum. Working paper. Tampere University of Technology. Virtual Building Laboratory, pp. 4. Available at: cic.vtt.fi/projects/vbe-

Hewahi, N. M., & Al Khateeb, Y. S. (2020, December). Neural networks representation for semantic networks. Int'l Conf. on Innovation and Intelligence for Informatics, Computing and Technologies (3ICT) (pp. 1-5). IEEE.

Horsbøl, A. (2018). Co-creating green transition: how municipality employees negotiate their professional identities as agents of citizen involvement in a cross-local setting. Environmental Communication, 12(5), 701-714.

Huemer, M. (1997). Probability and coherence justification. The Southern Journal of Philosophy, 35(4), 463-472.

Jiayi, P., Cheng, C. P. J., Lau, G. T., & Law, K. H. (2008). Utilizing statistical semantic similarity techniques for ontology mapping—With applications to AEC standard models. Tsinghua science and technology, 13(S1), 217-222.

Khalili, A., and Chua, D. K. (2013). IFC-based framework to move be- yond individual building elements toward configuring a higher level of prefabrication. Computing in Civil Engineering.

Kondyli, V., Bhatt, M., & Hartmann, T. (2018). Precedent-based design foundations for parametric design. Adv. Comput. Des, 3, 30.

Kraaijenbrick, J. (2007). Engineers and the Web: An analysis of real-life gaps in information usage. Inf. Process. Manage., 43(5), 1368–1382.

Laakso, M., & Kiviniemi, A. O. (2012). The IFC standard: A review of history, development, and standardization, information technology. ITcon, 17(9), 134-161.

Leake, D. B. (2001). Problem solving and reasoning: Case-based. In Smelser, N. J., & Baltes, P. B. (Eds.), International encyclopedia of the social & behavioral sciences (pp. 12117–12120). Pergamon.

Lee, J., & Kwan, M. P. (2005). A combinatorial data model for representing topological relations among 3D geographical features in micro-spatial environments. International Journal of Geographical Information Science, 19(10), 1039-1056.

Levitt, R. (2007) CEM Research for the Next 50 Years: Maximizing Economic, Environmental, and Societal Value of the Built Environment. Construction Engineering & Management, 133(1).

Lin, K., and Soibelman, L. (2009). Incorporating domain knowledge and information retrieval techniques to develop an architectural/ engineering/construction online product search engine. Computing in Civil Engineering. 23(4), 201–210.

Liao, L., Zhou, K., Fan, C., & Ma, Y. (2022). Evaluation of Complexity Issues in Building Information Modeling Diffusion Research. Sustainability, 14(5), 3005.

Linsbauer, L., Lopez-Herrejon, R. E., & Egyed, A. (2017). Variability extraction and modeling for product variants. Software & Systems Modeling, 16(4), 1179-1199.

Liu, X., Akinci, B., Bergés, M., & Garrett Jr, J. H. (2013). Extending the information delivery manual approach to identify information requirements for performance analysis of HVAC systems. Advanced Engineering Informatics, 27(4), 496-505.

Majumdar, A. K. & Sowa, J. F. (2009) Two paradigms are better than one, and multiple paradigms are even better, http://www.jfsowa.com/pubs/paradigm.pdf

Miettinen, R., & Paavola, S. (2014). Beyond the BIM utopia: Approaches to the development and implementation of building information modeling. Automation in construction, 43, 84-91.

Mortara, J., Collet, P., & Dery-Pinna, A. M. (2021, September). Visualization of Object-Oriented Variability Implementations as Cities. 9th IEEE Working Conference on Software Visualization (VISSOFT 2021).

Morganti, M. (2018). The structure of physical reality: Beyond foundationalism. Bliss & Priest 2018b, 254-272.

J. Miles, L. Moore, J. Cadogan, Matching computational strategies to task complexity and user requirements, Advanced Engineering Informatics 16 (2002) 41–52.

Munch, M., Buche, P., Dervaux, S., Dibie, J., Ibanescu, L., Manfredotti, C., ... & Angellier-Coussy, H. (2022). Combining ontology and probabilistic models for the design of bio-based product transformation processes. Expert Systems with Applications, 203, 117406.

Nguyen, T. H., Oloufa, A. A., and Nassar, K. (2005). Algorithms for automated deduction of topological information. Automation in Construction, 14(1), 59–70.

Nik Bakht, M., and El-Diraby, T. E. (2015). Synthesis of decision-making research in construction. Construction Engineering and Management, 141(9), 04015027.

Olsson, E. J. (2018). Bayesian Epistemology. In: Introduction to Formal Philosophy (pp. 431-442). Springer, Cham.

Palmer, S., & Udawatta, N. (2019). Characterising "Green Building" as a topic in Twitter. Construction Innovation.

Palmer, C., Harding, J. A., Swarnkar, R., Das, B. P., & Young, R. I. (2013). Generating rules from data mining for collaboration moderator services. Intelligent Manufacturing, 24(2), 313-330.

Petrinja, E., Stankovski, V., & Turk, Ž. (2007). A provenance data management system for improving the product modelling process. Automation in construction, 16(4), 485-497.

Pharo, N. (2008). "The effect of granularity and order in XML element retrieval." Inf. Process. Manage., 44(5), 1732–1740.

Poston, T. (2014). Reason and explanation: A defence of explanatory coherentism. Springer.

Pryke, S. D. (2004). Analyzing construction project coalitions: exploring the application of social network analysis. Construction Management and Economics, 22(8), 787-797.

Pryke, S. D. (2005). Towards a social network theory of project governance. Construction Management and Economics, 23(9), 927-939.

Pryke, S. (2017). Managing networks in project-based organisations. John Wiley & Sons.

Rabiser, R. (2019, September). Feature modeling vs. decision modeling: History, comparison and perspectives. 23rd International Systems and Software Product Line Conf. Volume B (pp. 134-136).

Rashid, Y., Rashid, A., Warraich, M. A., Sabir, S. S., & Waseem, A. (2019). Case study method: A step-by-step guide for business researchers. Qualitative Methods, 18, 1609406919862424.

Roche, W. (2013). Coherence and probability: A probabilistic account of coherence. In Coherence: Insights from philosophy, jurisprudence and artificial intelligence (pp. 59-91). Springer, Dordrecht.

Rogers, E. M. (2003). The Diffusion of Innovations, The Free Press, New York, 2003.

Salama, D. A., & El-Gohary, N. M. (2013). Automated compliance checking of construction operation plans using a deontology for the construction domain. Computing in Civil engineering, 27(6), 681-698.

Sowa, J. F. (2015). Why Has AI Failed? And How Can it Succeed? Available at http://www.jfsowa.com/pubs/micai.pdf

Surahyo, M., & El-Diraby, T. E. (2009). Schema for interoperable representation of environmental and social costs in highway construction. Construction Engineering and Management, 135(4), 254-266.

Tang, L., Zhang, Y., Dai, F., Yoon, Y., Song, Y., & Sharma, R. S. (2017). Social media data analytics for the US construction industry: Preliminary study on Twitter. Management in Engineering, 33(6), 04017038.

Tchomté, N. D., Asghar, S., Javaid, N., Dayang, P., Danga, D. E. H., & Oyono, D. L. B. (2020). A case-based reasoning coupling multi-criteria decision making with learning and optimization intelligence: Application to energy consumption. EAI Endorsed Transactions on Smart Cities, 4(9), e4-e4.

Thurimella, A. K., Bruegge, B., & Creighton, O. (2008). Identifying and exploiting the similarities between rationale management and variability management. 12th International Software Product Line Conf. (pp. 99-108). IEEE.

Turk, Ž. (2001). Phenomenological foundations of conceptual product modelling in architecture, engineering and construction. Artificial Intelligence in Engineering, 15(2), 83-92.

Turk, Ž. (2016). Ten questions concerning building information modelling. Building and Environment, 107, 274-284.

Turk, Ž. (2020). Interoperability in construction–Mission impossible? Developments in the Built Environment, 4, 100018.

Ueda, K. (2001). Synthesis and emergence—research overview. Artificial intelligence in engineering, 15(4), 321-327.

Van Berlo, L. A., & Natrop, M. (2015). BIM on the construction site: Providing hidden information on task-specific drawings. ITcon, 20(7), 97-106.

van Treeck, C., & Rank, E. (2007). Dimensional reduction of 3D building models using graph theory and its application in building energy simulation. Engineering with Computers, 23(2), 109-122.

Vanlande, R., Nicolle, C., & Cruz, C. (2008). IFC and building lifecycle management. Automation in Construction, 18(1), 70-78.

von Hippel, E. (1986). Lead users: a source of novel product concepts, Manag. Sci. 32 (7) (1986) 791–805.

Wang, H., Meng, X., & Zhu, X. (2022). Improving knowledge capture and retrieval in the BIM environment: Combining case-based reasoning and natural language processing. Automation in Construction, 139, 104317.

Wilde, N. (2020). Burning Bridges: The problem of relations in object-oriented ontology—a topological approach. Palgrave Communications, 6(1), 1-12.

Wix, J. (2007). Information delivery manual: Guide to components and development methods. BuildingSMART, 〈http://iug.buildingsmart .org/idms/development/IDMC_004_1_2.pdf/view〉

Yang, S., & González-Bailón, S. (2017). Semantic networks and applications in public opinion research. The Oxford handbook of political networks, 327-353.

Yurchyshyna, A., & Zarli, A. (2009). An ontology-based approach for formalisation and semantic organisation of conformance requirements in construction. Automation in Construction, 18(8), 1084-1098.

Zhang, J., & El-Diraby, T. E. (2012). Social semantic approach to support communication in AEC. Computing in Civil Engineering, 26(1), 90-104.

Zarei, B., Sharifi, H., & Chaghouee, Y. (2018). Delay causes analysis in complex construction projects: a Semantic Network Analysis approach. Production Planning & Control, 29(1), 29-40.

Zhang, L., & El-Gohary, N. M. (2016). Epistemology-based context-aware semantic model for sustainable construction practices. Construction Engineering and Management, 142(3), 04015084.

Zheng, X., Le, Y., Chan, A. P., Hu, Y., & Li, Y. (2016). Review of the application of social network analysis (SNA) in construction project management research. International J. of Project Management, 34(7), 1214-1225.

Zolin, R., Hinds, P. J., Fruchter, R., & Levitt, R. E. (2004). Interpersonal trust in cross-functional, geographically distributed work: A longitudinal study. Information and Organization, 14(1), 1-26.

APPENDIX I: Paradigms of knowledge representation

This appendix summarizes some of the theoretical and philosophical concepts related to the proposed approach. It provides a background about how we typically modeled AEC knowledge. It attempts to articulate a list of factors/issues that we should (re)consider while managing knowledge in the era of digital twinning.

The standardization-contextualization dilemma

Contextualization at the project or enterprise level is the first challenge to standards (Gogineni et al 2020). At the project level, context refers to the specific conditions/settings of a project. Contextualizing IFC is a key task for AEC practitioners. Research conducted with BIM users showed that they place an increased emphasis (and time) on understanding the context of information before using it (Hartman and Amor 2017). Some of the common factors that contribute to context uniqueness include the following:

The technical aspects of the project. This covers the type of project, the sophistication of the design and the construction approach.

(physical) Boundary conditions: these can range from natural attributes such as geotechnical, environmental and weather conditions to issues related to the local community and businesses, such as urban density and traffic volumes. Many of these concepts are ill-defined in their own domains and, in some cases, are even contested: what is sustainability anyway ? Maybe this is why a leading factor in derailing some (large-scale) projects is community objections (see Surahyo & El-Diraby 2009).

The business aspects of the project: Projects have *ad hoc* teams of stakeholders. They are always organized differently according to the project delivery system and dynamically throughout the project life cycle. This has a direct impact on their perspective on knowledge constructs and data models.

The modeling mentality

The most important paradigm of the proposed approach is that, without dismissing the value of standardized models, a no-model approach is valuable (indeed, needed) to supporting knowledge capture and information management in today's environment. One key drawback to adhering only to common models is their inability to capture uncertainty and change in knowledge constructs. Even projects with similar attributes repeated within similar contexts will have different knowledge constructs. This is, simply, because our knowledge evolves from project to project—we learn and technology progresses.

<u>Structuralism</u>: Increased penetration of BIM in the overall project information systems has exposed the significant limitations of the structured mentality (behind IFC): the inability to realize an artifactual system that achieves its purpose in unpredictable conditions (Kelly 2020).

<u>Emergence</u>: Traditional analytic approaches tend to be deterministic and based on top-down decomposition—for example, operational research, symbolic artificial intelligence, and knowledge-based engineering. In contrast, emergence-related approaches combine both the bottom-up and the top-down features—for example, evolutionary computation, behavior-based methods, reinforcement learning, and multi-agent systems (Ueda 2001). In computational emergence, at local levels, deterministic computational interactions scale to create global orders. This is referred to as self-reproduction or cellular automata. Of particular interest to this proposal is the concept of *relative-to-a-model*, which is fundamental to emergence research. Here '*difference*', captures emergence, which is measured in terms of deviations observed by a system user from expected or estimated behavior.

<u>Evolution</u>: while emergence refers to the recognition of system elements that were not necessarily recognized, evolution refers to the deviation of existing system elements from standard behavior. If we assume that we can formulate all aspects of emergence in knowledge constructs, and if we assume that we can address the contextualization problem, evolution will still be a problem for the standardized model approach. In time, the evolutionary nature of "human" knowledge will render some of these constructs and their connecting pathways irrelevant. The evolution of knowledge can be caused by innovations (in hardware or even project delivery systems) or, simply, learning new facts.

Epistemology

To handle the subjectivity and variability of AEC knowledge, many believed that a common model of knowledge is needed to streamline the modeling and exchange of data. Such a positivistic view has been shown to be inadequate because rigid conceptualization neglects to capture the (social) context of a project (Cerovesk 2011). Phenomenologists refute the existence of a model of truth; or our ability to capture reality. Models are subjective interpretations; their value stems from their appeal to more people, but not necessarily because they match reality (see Turk 2001). This gives rise to several considerations that should guide our modeling approach (see Cerovsek 2011): 1) Symbiosis of technology and technology use: we should learn from the practices of using standards and use that to guide the (re)modeling of our conceptualizations. 2) Knowledge is evolutionary: as practitioners continually attempt to improve, they create new practices and, accordingly, the structure of knowledge evolves; the models become more complex. 3) Crucially, recognizing the essential role of communication in creating knowledge; and the importance of (open) communication over the semiotics of communication (verbal and non-verbal between humans via computer).

<u>Ontological agency</u>: As an epistemology, constructivism (more accurately social constructivism) maintains that knowledge is a creation of its community which observes and seeks to create mental constructs to explain sensory experiences. Consequently, knowledge is a relative construct situated within the socio-cultural context of a community of practice; and is bound to evolve over time and place (Gergen, 1995). Two key points are relevant here:

Knowledge and power. According to Foucault, what we identify as knowledge is not (necessarily) a product of pure, systematic and impartial observation or analysis (by the community of practice). Common knowledge models can be a reflection of the views of the powerful stakeholders (in our case, experts who claim authority on domain knowledge).

Co-learning. How should the development and agreement on a knowledge construct happen? Should it be the result of purposeful top-down modeling by powerful experts? Or should it be automatically detected/generated by an algorithm? Or should it result from collaborative analysis, negotiation and co-learning between stakeholders? The latter assumes a knowledge agency for every stakeholder, emphasizes that knowledge(s) is multi-perspective, and empowers stakeholders to innovate.

New tools for bottom-up systems

A common belief behind BIM is that through developing and presenting exemplary cases and through the collective use of common standards, increased efficiency and economic gains will be realized—a common theme in normative thinking (see Murguia et al. 2021). This is a version of the classical theory of diffusion of innovation (Rogers 2003) and the theory of lead users in innovation (von Hippel 1986): the early adaptors set the course; the majority will learn from their experience and the overall system keeps moving to higher levels of optimality. Analysis of actual adoptions of BIM in AEC reveals that this type of adoption is only one of many (see Succar and Kassem, 2015; Liao et al. 2022).

<u>Synthesis vs analysis</u>: Analysis is an attempt to separate a whole into logical parts (decompose a structure to examine its functions). It aims at finding a principal theorization that captures our knowledge about complex systems; and then exposes the nature, function and interrelationships of its components. This approach has dominated IFC development. Synthesis, on the other hand, is a process of fitting a whole from parts or transferring individual propositions into complete systems. Analysis is better suited to clarifying causality (hence its consistent use in physics) and for reasoning (hence its appeal for decision support systems). synthesis research tools are more suitable for learning and discovering knowledge and innovating new worlds.

<u>Conformance vs difference</u>: with the model-based approach, we create a standard conceptualization for all to adhere to. However, if we aim to capture evolution, we should be looking for diversity in conceptualization, not conformance. We should detect what is new and what is different (in the exchanges of project teams); not suppress it. In fact, to capture the emergence of knowledge, we should challenge or deconstruct established norms (in conceptualization). This, as Derrida himself emphasized, should not be a deliberate technique because a thinker who adopts a purposeful method (of deconstruction) has already decided how to proceed. Rather, it should be about the discovery of emergent meanings (and contrasting them to established definitions).

Modal Logic: In propositional logic, validity can be defined using truth tables. A valid argument is simply one where true premises guarantee true conclusions. Modal logic relaxes the stringent inference rules of propositional logic. In modal logic, truth tables are not needed, because the semantics for modal logic can be defined by introducing possible worlds.

Prototype-oriented programming: Object-Oriented languages were built to use static type checking—the notion of a fixed class set at compile time. i.e., In these "class-based" systems, replication of classes happens at compile time. In IFC, all projects are to copy the standard classes in its scheme. Alternatively, in prototype-oriented programming (as in Javascript), the operations are stored in a prototype data structure, which is copied and modified at runtime. Should IFC continue to be an object-oriented model? Why should all concepts be locked in at the time of creating the standard? Can we allow users to clone classes with the ability to modify them? Can concept relations be contingent and editable? Can our systems update the classes (knowledge constructs) as we learn from the way users change the classes in practice? Ultimately, this is a question about what role precedence-based thinking plays in our work vs. model-based systems (Kondyli et al.2018).

APPENDIX II: A short list of example implementations of the proposed approach

The following table synthesizes a set of examples for implementing the proposed approach from previous research work. It includes eight cases. The table showcases that the source of data for establishing the networks can be through a survey of experts, text from documents, or social media posts or chats. Several methods are used to generate the networks from the data sources, including manual, semi-manual, and automated. Additionally, Boolean algebra can be used to artificially create new networks (worlds) from existing networks.

The objectives/outcomes of developing the networks can range from using expert opinion for building a baseline network that simply formalizes established knowledge of a group of experts (say within one company). This network can be used as a seed network for new projects. As the project knowledge constructs evolve, this generic network could also be used as a benchmark to study the uniqueness of the new project: How a new project is different from generic knowledge (as recognized within a company). Additional advantages of using the proposed approach and for the use of concept networks include formalizing project-level knowledge (a network of concepts vs. free text); tracking topic/concept morphism; prototyping constructs; studying actor network dynamics; examining levels of collaboration between actors; contrasting projects; and examine the creation of possible worlds.

Case #1: Developing a baseline network

For full details, see the paper by Aragao, R. R., and El-Diraby, T. E. on using network analytics to capture knowledge (Published in J. of Cleaner Production)

This case illustrates how to build a baseline network. This baseline network is to be updated frequently to act as the generic view of knowledge (say in a company or a domain). In this case, the research team conducted a literature review of the key concepts in one sub-domain: energy analysis/considerations during the design stage of Oil and Gas projects. The concepts were arranged into a matrix of relationships. A Software was used to create the network.

Case	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Main features								
Data source	Expert input	Text	Twitter	Twitter	Chats	Other networks	Other networks	Other networks
Methods for generating the network	Semi-automatic	Manual	Semi-automatic	Semi-automatic	Automatic	manual	mixed	Boolean algebra
Objective/Outcome	Build a baseline network	Formalize project knowledge	Community dynamics	Topic dynamics	Study collaboration	Contrasting projects	Prototyping	Examine possible worlds
Contribution	Capture base prototypes		Study knowledge dynamics			Benchmark-to-a-prototype		
Project-level analysis								
Social network analysis								
Qualify the network			X					
Study sub-communities			X					
Study influence			X		X			
The death of a network			X					
Semantic network analysis								
Identify topics	X	X		X	X			
Identify clusters	X	X		X				
Identify blocks							X	
Cross-project analysis								

Network-level analysis								
Network comparison						X		
Adding networks								X

Case #2: Developing a project-specific semantic network

For full details, see the paper by Nik Bakht, M., and El-Diraby, T. E. on social and semantic analysis of communities in infrastructure discussion networks (published in Computer-Aided Civil and Infrastructure Engineering).

And the paper by Aragao, R. R., and El-Diraby, T. E on using network analytics to capture knowledge (Published in J. of Cleaner Production).

The source of knowledge can be text reports developed by project participants (or social media posts). The research team manually transferred the text within the reports into a concept matrix and then a network. A set of terms (from the baseline network of Case#1) was used to build the two dimensions of the matrix.

Case #3: qualifying the network and studying its community dynamics

For full details, see the paper by Nik Bakht, M., and El-Diraby, T. E. on profiling Community Discussion Networks in Urban Infrastructure Projects (Published in J. of Infrastructure Systems).

This case illustrates how to handle and use social network analysis in the context of extracting knowledge constructs. This included the following tasks

Suitability of the network: using network analysis measures related to density and diameter to examine if the network portrays "small-world behavior"— basically that the social network is not a random graph and that it mimics a natural community of practitioners. The flip side of this analysis is the evaluation of the death of a network: when a network is no longer suitable as a source of representative knowledge.

Semantic network analysis: after understanding and qualifying the social network, using content analysis techniques to extract community interests and ideas discussed.

Network dynamics: monitoring the networks over time and studying how they evolve. In particular, the evolution of subcommunities and the migration of stakeholders from one to another as the project progresses is discussed; linking subgroups to specific topics.

Case #4: studying topic dynamics

For full details, see the paper by Nik Bakht, M., and El-Diraby, T. E. on project collective mind (published in Automation in Construction).

And the paper by Kinawy, S., Nik Bakht, M., and El-Diraby, T. E. on knowledge mismatches in stakeholder communication (Published in Sustainable Cities and Society).

This case tracked topics on Twitter accounts of major urban infrastructure projects. It showcases, the usability of the proposed approach with a larger dataset. The case integrated social and semantic analysis. To help measure public opinion, sentiment analysis was also conducted. Furthermore, the evolution of the networks was tracked as the project progressed from planning to construction to finishing.

Case 5: automatic generation of networks

For full details, see the paper by El-Diraby, T. E., Krijnen, T., and Papagelis, M. on the Green2.0 platform (Published in Automation in Construction).

Green 2.0 is a social BIM platform that represents a sample for applying the proposed approach at the project level. Users can comment and interact within the online BIM environment. Social and semantic networks were generated automatically. For each concept, sentiment analysis was also conducted. In addition to project-level network generation, a concept network was generated for each IFC product. Using network analysis measures, key indicators were developed and shared with the user in a dashboard.

Case #6: Contrasting projects

For full details, see the paper by Aragao, R. R., and El-Diraby, T. E. on network analytics and social BIM for managing project unstructured data (published in Automation in Construction).

This case showcases the use of the proposed approach for contrasting project-level networks. Five networks were available for analysis. A multitude of network analysis measures, particularly cluster analysis, were used to contrast the concept constructs in each project.

Case #7: prototyping constructs

For full details, see the paper by Aragao, R. R., and El-Diraby, T. E. on using Blockmodeling for capturing project knowledge constructs (Published in Advanced Engineering Informatics).

In this case, Blockmodeling was used to extract concept clusters from semantic networks. Blockmodels provide a more meaningful way of grouping the nodes (concepts) and finding their interrelations. This can be the basis for nominating knowledge constructs to become prototypes.

Case #8: creating future worlds

For full details, see the paper by Aragao, R. R., and El-Diraby, T. E. on using network analytics to capture knowledge (published in J. of Cleaner Production).

Using network algebra, several possible new worlds were created by adding existing networks from other/previous projects.