Towards a Digital Competency Framework for the Construction Industry

Mark Baldwin 1, Yiannis Xenidis 2, Fadi Castronovo 3, Sunil Suwal 4, Alexander Denzler 1, Thomas Heim 1, Angelo Luigi Camilo Ciribini 5
1 Lucerne University, Lucerne, Switzerland
2 Aristotle University of Thessaloniki, Thessaloniki, Greece
3 University of Brighton, Brighton, United Kingdom
4 Metropolia University of Applied Sciences, Helsinki, Finland
5 DICATAM University of Brescia, Brescia, Italy

Abstract
The construction industry is widely considered to be one of the least efficient, most wasteful and least digitalized sectors of the economy. Digitalization of the construction industry is acknowledged as a top priority, reflected in the massive demand for digital processes, as well as the increasing government mandates and regulations. Nevertheless, the lack of suitably skilled workforce and a standard mechanism for inferring digital competencies, means the industry is slow to reach digital maturity. This paper seeks to address this need, by proposing an architecture for a Digital Construction Competency Framework. This framework aims to support industry and academia in identifying necessary competencies for the digitalization of the construction industry.

Introduction
The European Commission’s newly formed High Level Construction Forum (HLCF) identifies three focus areas for the transformation of the construction ecosystem; “Green”, “Digital” and “Resilience” (referring to upskilling and industry competitiveness). The significance of these three distinct areas is summarized below:

- “Green”: The Construction industry accounts for 38% of CO2 emissions (UNEP 2019) in building and operation activities. Laden with inefficiencies and producing excess amounts of waste, the construction sector has experienced just a 1% annual increase in productivity over the past twenty years (Agarwal et al., 2016); therefore, a different model is urgently required to increase the industry’s efficiency, while reducing its carbon footprint.

- “Digital”: Digitalization, and especially “integrated BIM” (Building Information Modelling) is promoted as the most likely and most impactful technology towards transforming the construction industry (World Economic Forum 2016). The EU BIM Task recognizes the potential savings through digitalization in Europe alone as being in the range of EUR 130 billion per year (EU BIM, 2018).

- “Resilience”: In the EU alone, the construction sector accounts for 18 million jobs and contributes to almost 9% of the GDP (European Commission 2020a). The report by Arden University on the effect of automation on the UK’s working environment in various sectors revealed that a 20% of the construction sector’s employees feel unprepared and inadequately skilled to deal with the industry’s digitalization demands (Arden University 2022). A quite similar situation is met in the European construction sector, where “skill mismatch” is reported to be the most important concern for a 79% of the construction companies operating in the EU (European Construction Sector Observatory 2021). Another important finding of the ECSO (2021) report is that in most European countries the significant increase of unfilled vacancies related to Information and Communication Technologies (ICT) in the construction sector has not been followed by a respective increase in the provision of training. In other words, the digital gap in the construction industry is not sufficiently treated as training and education on digital competencies have not grown in line with the industry’s demands (European Construction Sector Observatory 2021).

Being a key driver for the economy worldwide, it is evident that the longer it takes for the construction sector to transform to digital ways of working, the less any sustainability and environmental strategies will be effectively implemented. This slow transformation is the reason behind the many and various types of efforts that seek to expedite the digitalization of the construction industry. Policy initiatives include the “Strategy for the sustainable competitiveness of the construction sector and its enterprises” (European Commission 2012), the “Renovation Wave for Europe” (European Commission 2020), and the ‘new Circular Economy Action Plan’ (European Commission 2020).

Another tool for achieving the required transformation is standardization. Several national standards focusing specifically on Building Information Modeling (BIM) competency definition have been published, including the Italian UNI 11337-7 and the German VDI 2552-8.2. Standardization is necessary to enable the successful implementation of digital processes, however, standards are largely retrospective in nature, documenting the
current state of the art, and not necessarily accounting for future innovation and development. (Cargill, 2011).

The European Committee for Standardization, CEN, formed a technical committee (TC 442) to address the standardization of BIM. In 2021 a new work group of TC442 (WG8) was launched to address the issue of digital competences around digital ways of working. An initial finding of this work group is that current national standards that have been published in this domain tend to be narrow in scope. This means that the BIM tasks and competences defined in one country are not necessarily applicable in another country. Indeed, as digital processes evolve these standards become quickly outdated. A proposal put forward by the WG8 is that a general framework could be developed which outlines general principles for defining digital competences. The output would not be a standard, but rather a technical report or specification.

An alternative to a top-down standards approach could be an industry-lead mechanism for defining best practice workflows and the inferred digital competences.

This paper proposes such a technology driven, bottom-up approach within the concept of a Digital Construction Competency Profiler (DCCP). The proposal is to define processes for digital ways of working in a standardized way; to infer the necessary competencies to fulfill the given tasks; and to provide a mechanism for establishing custom profiles and identifying candidates ideally suited to fill those profiles.

DCCP is a joint effort between academia and research institutions (Lucerne University of Applied Sciences and Arts in Switzerland - HSLU, the Swiss Federal University - ETH and the Fraunhofer Institute) and construction industry associations (the Swiss Centre for Standardization in Construction - CRB, Swiss Association of General Contractor - SBV, and the European Council for Computing in Construction - EC3) that aims at fostering the achievement of a more skilled and competitive workforce and refining the focus as well as improving the quality of education and training in the construction sector.

**The Digital Construction Competency Profiler**

The main idea behind DCCP is that it builds on accumulated knowledge to define the required competencies per case of application, instead of defining a knowledge and skills background that potential employees should respond to. As shown in Figure 1, this approach leads to potentially deriving new best practices (thus innovating at the same time), instead of applying the best practices included in a prescriptive standard (thus restricting applications to the state-of-the-art). The following subsections present the architecture of DCCP structured step by step.

**Process Overview**

The research work begins by identifying relevant concepts from different sources, such as job description, CVs, education curricula and industry competency standards (Fig.2 Stage 1) and establishing relationships between these concepts (Fig.2 Stage 2). Some preliminary work has been undertaken to compile a list of established terms in this domain, as defined by various ISO standards (refer to following subsection). Preliminary manual work will be augmented by automated processes, namely using NLP and statistical analysis to extract new concepts and establish more domain specific relationships from the industry data sources.

These concepts are logically arranged according to expert inputs into an ontology, that we call the “competency framework” (Fig.2 Stage 3). Preliminary work has also been undertaken here to establish a base structure of the competency framework (described later in more detail).

**Figure 1: Top-down Vs Bottom-up approaches for competencies definition**

**Figure 2: Components of the Digital Construction Competency Profiler**
On-going data extraction and data validation will continuously refine the framework and gradually populate it with data to create a domain specific Knowledge Base (Fig.2 Stage 4) domain specific knowledge base. The combination of NLP-based automation and expert input, in addition to competence standards from the construction industry provide a novel and unique basis for further inference. This includes being able to define competency requirements for new job profiles, vocational planning or development of new education programs, all fully data driven and based on the identified industry insights Fig.2 Stages 5 and 6).

The Competency Framework

The first step is the building of an ontology that is called the Competency Framework (CF) and defines the relationships of the DCCP’s core concepts. Building on established terms in this domain, the CF will create a catalogue of concepts, and map relationships between these concepts. The following established terms are given as an example:

- “Actor” (according to ISO 29481-1:2016, part 3.1): a person, organization, or organizational unit (such as a department, team, etc.) involved in a construction process.
- “Certificate” (according to ISO/IEC 17024:2012, part 3.5): a document issued by a certification body under the provisions of this International Standard, indicating that the named person has fulfilled the certification requirements.
- “Competence” (according to ISO/IEC 17024:2012, part 3.6): the ability to apply knowledge and skills to achieve intended results.
- “Continuing Professional Development (CPD)” (according to ISO/IEC TS 17027:2014, part 2.27): the activities undertaken by a person after initial education or training to maintain, improve or increase his/her knowledge and skills related to his/her professional activities.
- “Knowledge” (according to ISO/IEC TS 17027:2014, part 2.56): the facts, information, truths, principles or understanding acquired through experience or education.
- “Qualification” (according to ISO/IEC 17024:2012, Part 3.7): a demonstrated education, training, and work experience, where applicable.
- “Skill” (according to ISO/IEC TS 17027:2014, part 2.74): the ability to perform a task or activity with a specific intended outcome acquired through education, training, experience, or other means.
- “Training” (according to ISO/IEC 17024:2012, Part 2.77): a programme developed to provide persons with the necessary knowledge and skills.

The rudimentary schema provided in Figures 4, 5 and 6 show a possible relationship of these concepts and how they could be assembled to infer competences from tasks. The fundamental premise in the ontology’s structure is that a profile is defined by a set of tasks associated with a use case. Each task requires specific skills, knowledge and behavior (i.e., competencies) in order to be fulfilled; however, tasks, and the competencies needed to fill them, vary depending on contexts such as the geographic location, industry sector, project type, project phase, etc. Moreover, a client may identify additional requirements, such as the use of specific software or certain industry qualification, necessary for the fulfilling of the proposed role. Therefore, these components, grouped as shown in Figure 3, altogether determine the profile characteristics.

Each component’s element or concept may have multiple constituent parameters, and each parameter can have multiple values. For example, the concept of Education is specified by the Field of Study (e.g., Architecture or Engineering), the Level of Study (e.g., Diploma, Bachelor, MSc, PhD), the Duration (years) of studies and perhaps also the Date of Completion. Similarly the concept of Knowledge can be subdivided into the following parameters and values: Knowledge set (content), Depth of knowledge (evaluated with a rating scale of 1-5 based on experience), and Application (evaluated, again, with a rating scale of 1-5 based on experience).

Figure 4 and Figure 5 depict two indicative deployments of specific entities in the ontology, namely the Profile Definition and the Additional Requirements. The presented entities were selected randomly as the detailed illustration of all entities was not practical due to space limitations.

Figure 5 illustrates the ontology of the Competency Framework. In this figure, a critical component is the “use-case” entity, which is the critical element for defining a profile, identifying the tasks to be undertaken and the actors required to fulfill the tasks. To facilitate the accurate description and proper introduction of use cases in the Competency Framework the Use Case Management (UCM) service developed by buildingSMART International is exploited (https://ucm.buildingsmart.org/).
UCM is a bottom-up approach that enables individual users and user groups to specify and exchange their BIM best practice use cases, (i.e., digital processes) in a standardized machine-readable format, based on the IDM Methodology (ISO 29481-1 & 3). UCM is an open platform, driven by users and fostering a collaborative way to define and standardize best practices of the future. The platform has gained widespread acceptance in the global BIM community as a quasi-standard for defining digital ways of working. It already contains over 100 well defined use-case definitions and many leading owner organizations, such as Swiss Rail, are using the UCM platform to define their BIM use-cases. However, while the platform can be used for presenting and exchanging best practices, it cannot be used for deriving competencies. This functionality is provided by the DCCP with the use of the Knowledge Base component.

The Knowledge Base

A core activity of this proposal is the creation of DCCP’s Knowledge Base (KB), that will be exploitable for diverse uses, such as the definition of competency requirements for new job profiles, vocational planning or development of new education and training programs, etc. The Knowledge Base will evolve progressively through a continuous process of data accumulation and exploitation. More specifically, large data sets will be extracted from use-cases, professional and academic sources (e.g., job descriptions, education curricula, competency standards, etc.) and classified according to the Competency Framework. Using AI supported data processing, such as NLP and statistical analysis, new concepts will be derived and more complex relationships between the ontology’s parameters will be identified.

As this activity is ongoing, the Knowledge Base will be continually grow, updated and expanded. Ultimately, this could become the most comprehensive, machine-readable database of digital competencies for the construction industry.

Application of Natural Language Processing (NLP)

The relevant scientific and technological knowledge for this project is multifaceted as follows.

The first step is to extract the concepts of interest, adopted to computationally extract collection of synonyms, extracting their relevance for the given context of application [Ventura 2015]. The resulting computational space for characterizing natural language artifacts is normally based on shallow Neural Networks [Waldis 2018] and uses multiple extraction techniques, such as ESA [Mazzola 2020] or WORD2VEC [Mikolov 2013]. The main limitation existing on this step is the absence of a specialized controlled vocabulary for the digital construction domain, that will be required to represent skills and tasks in this domain. Secondly, the relatedness of the concepts plays a fundamental role, allowing to move from isolated nodes to a network-oriented structure. In this context, we will explore different variations to find a suitable solution, while testing the achieved effectiveness. To make the Concept Map usable, a hierarchical organization within the concepts set is the next natural step. This can be obtained by clustering together concepts that present a minimum level of connectivity, also using fuzzy approaches [Pedrycz 2014] to consider the uncertainty and the variations present in real-world applicable ontologies.

On top of clusters identification their organization into hierarchical levels is important, following the insights from [Denzler 2019] and in compliance with the field of granular computing [Zahed 79] with the objective of embedding a semantic meaning natural for the given domain. This allows to use the concepts of abstraction and specialization to reason over them and thus to understand specificity and coverage provided by every set of concepts grouped into a cluster, called granule. Anyway, as demonstrated in [Stalder 2021], there is no current generally accepted general approach for achieving the granulation, but many variations that suit one or more use cases, thus our contribution in this respect will be to clearly understand the specific requirements for a digital
Construction-specialized granular knowledge base (GKB) and to develop a standardized and adaptable approach for the knowledge structuring, using granular computing. Eventually, granules automatically extracted from natural languages processes can be useful to understand the specific domain knowledge, but have limitations being purely based on statistical and distributional properties of the languages. Multiple tries exist in literature to achieve this automatically using graph structures [Kiwimaki 2013] or for specialized tasks, such as education path suggestions [Vo 2022] or for curricula analytics [Kitto 2020].

Figure 6 shows the hierarchical, granular structure of the knowledge base with separate layers for users, use cases and projects. The structure of the technology makes it possible that this data can be analyzed for innumerable purposes: for example, to determine the competency differences, as well as salary differences, for the same job in different regions, or indeed in different sectors in the same market. We will also be able to extrapolate what are the skills that are in most demand, by role, by market sector or by region. Thereby having a solid basis for making forecasts about the education needs and focus areas of the future. Achieving this level of granularity has not been attempted in this domain and the benefits will be far reaching.

**Validation**

The main innovation of this project lies in the creation of a domain specific knowledge base that manages to structure data in a hierarchical way on multiple levels, based on their level of granularity. As a result, novel insights can be extracted such as the complexity by simply looking at the average distribution of concepts within the levels. From the fuzzy affiliation of concepts to clusters, we can infer much richer recommendations, as we would from a flat affiliation. In addition to that, the separate layers for users, use cases and projects that are interconnected through the knowledge base, allow us to create custom profiles for each layer. This is a novel setting, in combination with the hierarchical structure, unlocks the full potential. Leading to even better and more accurate match making, finer grained selection of relevant Top-N results and the ability to personalize the result computation with the use of weights.

In order to favour expert knowledge to be explicitly added and combined with the GKB, we will allow human specialists from architecture, engineering and construction to specify skills as highly meaningful characteristics to describe tasks, job openings, candidate profiles and educational paths. This new view will be interconnected with the semi-automatically computed NLP pipeline through handcrafted relationship types and they will create an additional layer of the knowledge base, thus providing a validated view over the computed GKB. Another relevant aspect is the users secure identification and the validation of the information provided by applicants in the systems, this will require a Self-Sovereign Identity (SSI) solution [VanBokkem 2019] and the identification of a secure still scalable and sustainable approach to data vetting on a public ledger [Dubovitskaya 2019].

**Case study: Use case for cost estimation**

Use-cases defined according to the ISO 29481-3 Standard (IDM3 – XML Schema) have a standard structure and are machine readable. As such they can be computationally assimilated into the DCCP’s Knowledge Base. Figures 7a and 7b present an example of free text and images that could be used for a cost estimation use case, while Figure...
8 presents the respective structured process-map that identifies, which actors perform what tasks at specific phases in the project.

Use cases are embedded with meta data that can, for example, identify the project’s geographic location (country), as well as the project’s phase, type, etc. With this meta data it is possible to extrapolate the difference between two comparable use cases originating from two different regions or associated to different project types.

In the context of the cost estimation use case, a competency statement from the Royal Institute of Charter Surveyors (RICS) in United Kingdom is analyzed using NLP. This industry standard provides a detailed overview (from a British context) of the tasks, skills knowledge and behavior that is expected from a quantity surveyor (or cost estimator).

The use case of cost estimation could also include data from job descriptions from another region, e.g., the US as shown in Figure 10. Such data provide not only a detailing of the tasks and expected competencies from the point of view of the employer, but also identify minimum education requirements and even suggest a salary range.
Such data when integrated in the Knowledge Base provide with context specific particulars that would be very useful for several cases, such as recruitment, market and sector analysis, education and training programs planning, etc.

Conclusions

The construction industry faces a critical deficiency; namely, the lack of a comprehensive mechanism for defining digital competencies across the global construction sector.

This paper presents the concept and architecture of a new technology infrastructure, the Digital Construction Competency Profiler (DCCP), that will enable competency profile creation and matching with unparalleled speed and accuracy. DCCP has the potential to become a game changer in enabling professionals, companies and/or training providers to precisely define their competency profile requirements. This will lead to more effective matching of professionals with potential employees, service providers and/or educational services; ultimately accelerating the sector’s digitization.

DCCP through a dynamic evolution process can reach high levels of granularity in defining competency profiles and conversely in the inference algorithms for testing candidates against these profiles.

All data extracted is laden with meta data relating to its origins, for example its geographic location, whether it relates to a specific construction sector (e.g., general construction projects infrastructure), a specific project type and/or specific project phase. So, for example, it is possible to identify granular differences of a generic task such as “cost estimation”, depending on whether this is assigned (i.e., in China or Chile); whether the project may be a bridge or an airport and what phase the task should be performed in. This data can be analyzed for innumerable purposes, such as the determination of competency differences, as well as salary differences, for the same job in different regions, or indeed in different sectors in the same market. Another possibility is the extrapolation of skills that are in most demand, by role, by market sector or by region, thus acquiring a solid basis for making forecasts about the education needs and focus areas of the future.

Acknowledgments

Acknowledgement is also given to the preliminary research undertaken by the DCCP project team at Lucerne University of Applied Sciences and Arts and Bimondis.

References


