



**HAL**  
open science

## From digital traces to competences

Kamel-Eddin Bemmami, Lilia Gzara, Jean-Luc Maire, Christophe Courtin

► **To cite this version:**

Kamel-Eddin Bemmami, Lilia Gzara, Jean-Luc Maire, Christophe Courtin. From digital traces to competences. IFAC-PapersOnLine, 2022, 55 (10), pp.1944-1949. 10.1016/j.ifacol.2022.09.683. hal-03832624

**HAL Id: hal-03832624**

**<https://hal.science/hal-03832624v1>**

Submitted on 27 Oct 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

## From digital traces to competences

K-E. BEMMAMI\* L. GZARA\*\* J-L. MAIRE\* C. COURTIN\*

\*Univ. Savoie Mont Blanc, SYMME Annecy, France

\*\*INSA Lyon, DISP Lyon, France

*kamel-eddin.bemmami@univ-smb.fr*

*lilia.gzara@insa-lyon.fr*

*jean-luc.maire@univ-smb.fr*

*christophe.courtin@univ-smb.fr*

**Abstract:** Many studies have highlighted the fundamental role of human capital in the inventory of an organization's values, and more particularly in its positioning in relation to its competitors. For an organization, a thorough knowledge of people and their competences allows it to have good visibility of its internal resources and to build a sustainable competitive advantage. With the digital transformation of organizations and increasing complexity in business processes, actors would learn a lot and redefine continuously their working practices. Competence management has become an essential factor in industrial processes' performance. However, competences management requires a relevant system of competence identification, given its complex nature and the evolution of the company's activities. In this paper, we propose an activity trace-based system to identify and represent competences mobilized during an activity. A specification of terminology is first defined around the concept of competence and digital trace. This is followed by the presentation of the designed system and the result of its applicability on an industrial case study.

Copyright © 2022 The Authors. This is an open access article under the CC BY-NC-ND license (<https://creativecommons.org/licenses/by-nc-nd/4.0/>)

**Keywords:** competence formalization, digital traces, Enterprise information systems, human resource management.

### 1. INTRODUCTION

With the digital transformation of organizations in the industry 4.0 context, industrial work (design, production, distribution) is profoundly renewed. Human-machine interfaces will be at the heart of the issues. The major challenges of this revolution are not only technological but also, in the same way, related to human capital (Pinzone et al., 2017). As digital technologies become commodities in manufacturing operations, success depends largely on organizations and their employees. The human operator is confronted with the increasing complexity of his daily tasks becoming relatively less repetitive and requiring ongoing competence development (SIMIC and NEDELKO, 2019). However, the extent to which human competences evolve in response to changes in manufacturing technology has a direct impact on the sustainability of companies (Hecklau et al., 2016).

To overcome the challenge that represents the human capital, we need first to identify and formalize the actual competences mobilized by the actors across daily activities. Secondly, an adapted representation is needed to optimize the exploitation of the acquired knowledge for improved management of human resources and enterprise performance. These two points constitute the core of this paper which proposes a system of identification, formalization, and representation of the competences effectively mobilized in companies.

Because of the limited operational scope of the concept of competence, which has been observed in the practices of organizations, we focus on an in-depth analysis of the concept of competence (§2). We then present an overview of the trace concept and its use in terms of competence identification (§3). Our activity trace-based system for competence identification is then described (§4). We conclude with a presentation of the application of our system to an industrial case study (§5).

### 2. COMPETENCE

The concept of competence can be traced back 3000 years when the Chinese empire replaced the system of recommendation for high state positions with a system of examinations of the skills necessary for the proper performance of the job (Wilcox and King, 2013). However, the term 'competence' was first used in 1959 by White in his paper 'Motivation Reconsideration: The concept of competence' (White, 1959). Since then, the way of understanding a competence has changed a lot. Table 1 lists some recent contributions to the concept of competence.

**Table 1. The concept of competence**

<b>El Falaki, Idrissi and Bennani, 2010</b>	Competence organization as a relation linking three areas, firstly, Knowledge described by concepts & procedures. Secondly, Skills: describe the processes that can be applied to domain knowledge. Thirdly, public target: describe behavioral skills, characteristics, functions, and tasks or factors of competence.
---------------------------------------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

<b>Robert, 2012</b>	Competencies include the collection of success factors necessary for achieving important results in a specific job or work role in an organization. Success factors are combinations of knowledge, skills, and abilities.
<b>Kou, Jia and Wang, 2013</b>	A competence emphasizes the connection between the holistic complex of knowledge, skill and motive, attitude, etc., and a job fit under the specific work context.
<b>Kolibáčová, 2014</b>	The concept of competence is usually applied to define the whole of individual abilities, skills, behaviors, and knowledge, oriented to effective performance in a particular working environment.
<b>Chouhan and Srivastava, 2014</b>	Competence is the capability of applying or using knowledge, skills, abilities, behaviors, and personal characteristics to successfully perform critical work tasks, specific functions, or operate in a given role or position.
<b>Miranda et al., 2017</b>	The ability of an actor to do actions for a situation effectively and efficiently; it cannot be directly measurable but estimated from the performance.

Considering the different definitions mentioned above, we can observe that contrary to a knowledge that rather applies to theoretical aspects, competence also applies to practical aspects. In fact, competence refers to the use of knowledge in a specific situation.

In our works, we define competence as a capability of an actor to mobilize material and immaterial resources to perform a finalized activity in a situation that leads to superior performance toward customer satisfaction.

### 3. TRACE

The term trace is widely used and seems to be easy to define. But this is unfortunately not really the case. In the scientific literature, many research works deal with the notion of trace, but there is no standard definition to this day. Two sources of trace can be distinguished when dealing with activity trace. First, traces resulting from human observation and activity analysis. Second, traces resulting from digital tools used by actors when performing activities. In our work, we focus on digital traces and consider the concept of trace in terms of activities<sup>1</sup> evidence left in digital tools. Indeed, as many researchers state, industrial activities could not be run without producing digital traces of their execution (Hand and Gorea, 2018). These traces might be e-mails, written documents, logs, tags visible in DWE (Digital Working Environment) but also metadata on the time spent during an activity, on allocation or attribution of resources, etc. Table 2 lists some recent contributions to the concept of digital trace.

**Table 2. Authors contributions to the concept of digital traces**

Auteurs	Definition of digital traces
<b>Courtin and Talbot, 2005</b>	A digital trace is a signal which corresponds to a time pinpoint and elementary element containing several attributes (source, date, event description, parameters), or a sequence that splits into a chronological succession of signals or sub-sequences.
<b>Settoui et al., 2009</b>	Trace of the activity of a user who uses a tool to carry out this activity saved on a numerical medium.
<b>Lund and Mille, 2009</b>	A temporally situated sequence of observations, which is either an interaction between humans, mediated in various ways by a computer or a sequence of actions and reactions between a human and a computer

<sup>1</sup> Activity: a sequence of actions carried out in order to achieve a clearly defined objective (Belkadi, Bonjour and Dulmet, 2007)

<b>Zarka et al., 2012</b>	A record of the actions performed by a user on a system, in other words, a trace is a story of the user's actions, step by step
<b>Alecu, 2013</b>	Digital footprints (also known as digital traces) are the marks a user had left by the interactions in a digital environment. These marks are usually reflecting details about the user's activity in the digital environment (IP address, items clicked, elements searched for, visited pages, calls content...).
<b>Wang, 2016</b>	As a record of the user's interactive activity within the framework of the system. This interactive activity happens between users and has an objective to reach.

Considering the various definitions in Table 2, we use the definition of (Settoui et al., 2009) in our works. We also consider digital traces to be dynamic as they are a record of actions performed on a computer, Web solution, software, or any other digital interface

### 4. FROM DIGITAL TRACE TO COMPETENCE

Extracting knowledge from digital data is now a major challenge for the competitiveness of organizations (Coron, 2018). Computational processing and retrieval techniques have evolved considerably in recent years (Belmonte and Segura-Robles, 2020). As a result, some works propose automated decision-making systems based on employees' performance and productivity assessment data (Nicolaescu et al., 2020). With regard to the concept of competence, research projects involving digital traces can be categorized into two groups:

- competences inventory: research on industry 4.0 and human resource management has led to the creation of relatively structured competence repositories, based on statements of need or supply (Hecklau et al., 2016; Graczyk-Kucharska et al., 2018).
- competence-gap identification: numerous works on competences and exploiting digital traces aim at bridging the gap between the competences expected by the employment market and the CVs of candidates (Miranda et al., 2017).

In both cases (whether in job offers, candidates' CVs, or annual interviews conducted by HR department), considered competences are declared by people or retrieved from competences repositories. In order to reduce errors related to human subjectivity when declaring competences and risks of forgetting some of them, we propose in this study to extract competences from actual work practice. To that end, we propose a system to identify competences from digital activity traces. Figure 1 shows the architecture of the proposed system.

Our system is composed of four parts: trace generator, data processing, performance calculator and competence visualizer. In the following, we describe each of these parts.

#### 4.1 Trace generator

Digital activity traces may have three forms: logs, tags, and textual elements. Logs correspond to documentation of events concerning a particular system that is automatically generated and time stamped. Virtually all systems and software produce

log files. Tags are elements of information (such as an Internet bookmark, database record, or computer file). This kind of metadata helps describe an item and allows it to be found again by browsing or searching. Textual elements are any written production generated through a digital solution (e.g. emails, PDFs, electronic books...).

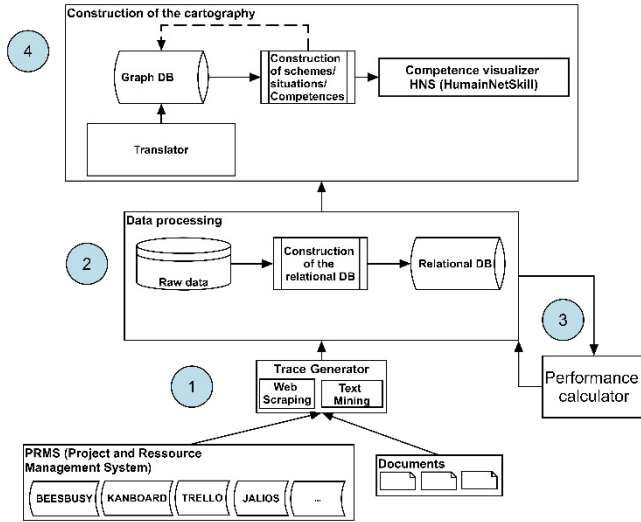


Figure 1. Competence identification system architecture

Logs' exploitation requires a full access to the digital tool core in order to extract logs. However, most companies use commercial tools and hence the access to logs remains limited. Moreover, logs' content are more often very limited in terms of information nature (e.g. activity, resources, actors...). For all these reasons, we have chosen to base our system for competences extraction on documents and tags.

Concerning, tags, they come from project and resource management systems (PRMS). Trace generator uses Web scraping technics to extract data from tags. As shown in Figure 2, the scraping allows us to recover data (e.g. the heading, the state, the estimated time, the actual time...) on projects, activities, actions, and resources, in a structured way.

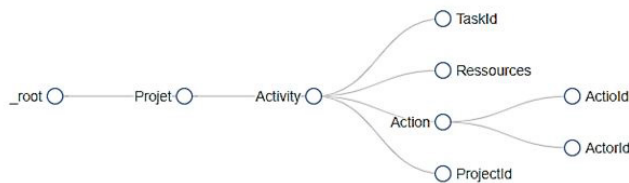


Figure 2. Scraping map

Concerning textual elements, the trace generator uses text mining algorithms to recover data. Text mining follows two main steps as presented in Figure 3. The first consists of analysing the corpus of texts in order to recognise the words, sentences, grammatical roles and the relations and meanings between them. This first stage, common to all treatments, is only relevant when coupled with the second stage: the interpretation of data. This step allows the selection and classification of particular texts among others using a semantic query.

In our approach, we assume that data generated from the Web scraping technics and text mining algorithms are complementary.

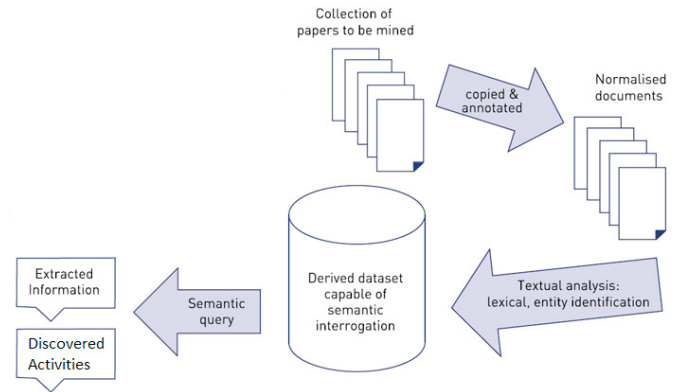


Figure 3. Text mining process

#### 4.2 Data processing

Even if the data recovered through the trace generator (through Web scraping or text mining) are structured (e.g. activity + objective, resources + actors ...), they are presented as a single block. In this format, the data is difficult to manipulate and the links between the various elements are absent. These data need to be processed and added to a database. The transition to a relational database allows the use of the SQL language to manipulate data more easily. Data processing is therefore the second step of our system.

Either in text mining or Web scraping, data obtained through trace generator are often exported in the structured CSV format. Before integrating data into the database, we first import the CSV into an Excel sheet, which will allow the use of specifically developed macros. These Excel macros allow processing of data in order to extract information on projects, activities, objectives, resources, actors, actions, and the links between these elements.

Figure 4 shows the data model we use.

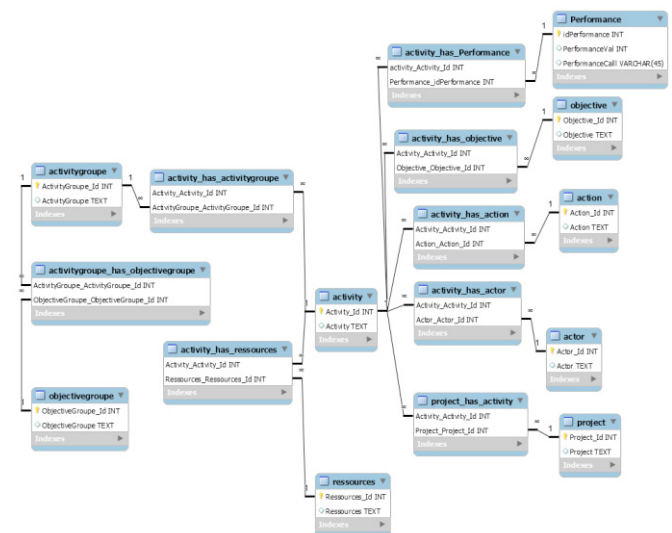


Figure 4. Relational database model

### 4.3 Performance calculator

The performance calculator is the third part of our system. The performance of each actor's activity is assessed based on two indicators: effectiveness and efficiency and under the assumption that the resources were relevant to achieve the objective.

The effectiveness (F) is evidenced by the state of the activity in PRMS. The fact that an activity is assessed as achieved indicates that the objective is fully or partially achieved. The effectiveness of activity can only be judged once the activity has been completed (ongoing activities are not evaluated).

Each activity in PRMS contains information on the estimated time and the actual time. The difference between both gives information on the level of efficiency. The efficiency (E) is considered on the base of five degrees scale (1: Not efficient / 2: Somewhat efficient / 3: Efficient / 4: Very efficient / 5: Extremely efficient). The interval of time for each level is defined in agreement with the project owner so that the information deduced is as close as possible to reality.

The performance (P) of the actor (A) on the activity (C) is then calculated according to the following formula:

$$P(A)_C = E * F, \quad E \in \{1, 2, 3, 4, 5\} \text{ and } F \in \{0, 1\}$$

If the performance value is 0, no information can be given because the activity has not yet been carried out (in this case, the effectiveness cannot be judged). If the value is different from 0, the result will be interpreted as follows: 1: Not performant, 2: Quite performant, 3: Performant, 4: Very high performance, 5: Extremely high performance.

The information on the performance is then stored in the relational database.

### 4.4 Building of the cartography

Graphical visualization of data helps to identify quickly important elements and the correlations between them. In this sense, it helps users to exploit competences related data in a graphical representation which is easy to understand. The use of graph theory for more advanced analysis is also an important point that leads us to choose a graphical representation for the visualization of competences.

For the construction of a graphical view for our data, we need to transform the relational database into a graph database. Some translators allow an easy transition to a graph database, such as the translator proposed by Neo4j<sup>2</sup>. This cartography construction phase allows us to use Cypher<sup>3</sup> queries in order to identify and represent for each activity: the actions needed and the situations in which it occurs. We assume that the situation of an activity includes the project to which it is attached, the activity group to which it belongs and the used resources.

With the set of elements identified and represented, we can identify the potential competences, by interacting with the graphical database. To interact more easily with the graphical

database, we have programmed queries in a graphical interface that we named 'Competence visualizer'. Across this visualizer, it will be easier to analyze the competences of the organization, the most frequent situations, and the collaboration between the actors across the shared activities. The view on competences will also allow identifying the critical competences that are owned by a unique actor and the activities that are related. Figure 5 shows the meta-graph for competence representation in the visualizer.

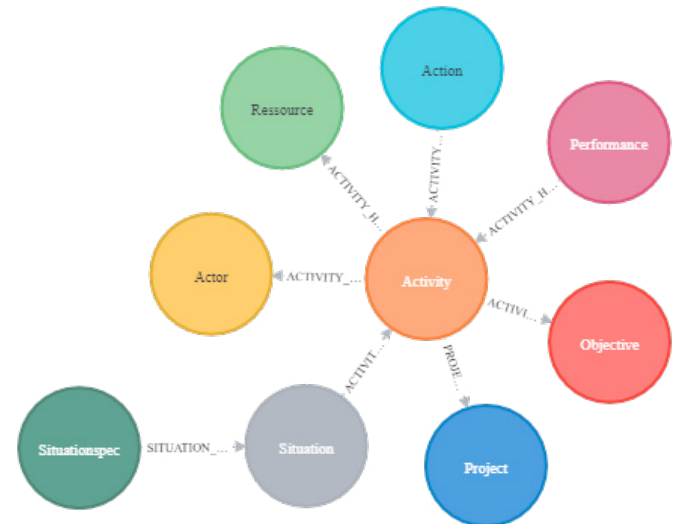


Figure 5. Competence meta-graph

The oriented link (arrows) between nodes represents the influence of element A (arrow origin) on element B (arrow target).

## 5. EXPERIMENTATION

Our system has been used and applied in Manaslu Ing<sup>4</sup> which is an engineering and consulting company with strong technical expertise in the field of buildings' energy efficiency.

This company uses Beesbusy to manage its projects (PRMS Web solution).

For the case study objectives, company employees were asked to follow a reporting formalism when declaring their activities on Beesbusy tool. The formalism includes a structure for the description of the activity in the form of "Activity verb in infinitive" + "for/ in order to" + "Activity objective". This statement structure allows the activity to be best described in relation to its objective. Actors were also asked at the end of each activity to specify the mobilized resources and the actions that were carried out. Beesbusy offers the possibility to declare the actions in a dedicated field, however, for the resources we asked that it be filled in a comment.

The experimentation was conducted for two months during which 132 activities were declared on Beesbusy by two employees. The use of Web scraping technics to generate digital traces from Beesbusy solution resulted in the recovery

<sup>2</sup> Neo4j is an open-source graph-based database management system. <https://neo4j.com>

<sup>3</sup> Cypher is a graph-oriented query language used by Neo4j. <https://neo4j.com/developer/cypher/>

<sup>4</sup> Manaslu Ing: <http://www.cmdl.fr/>

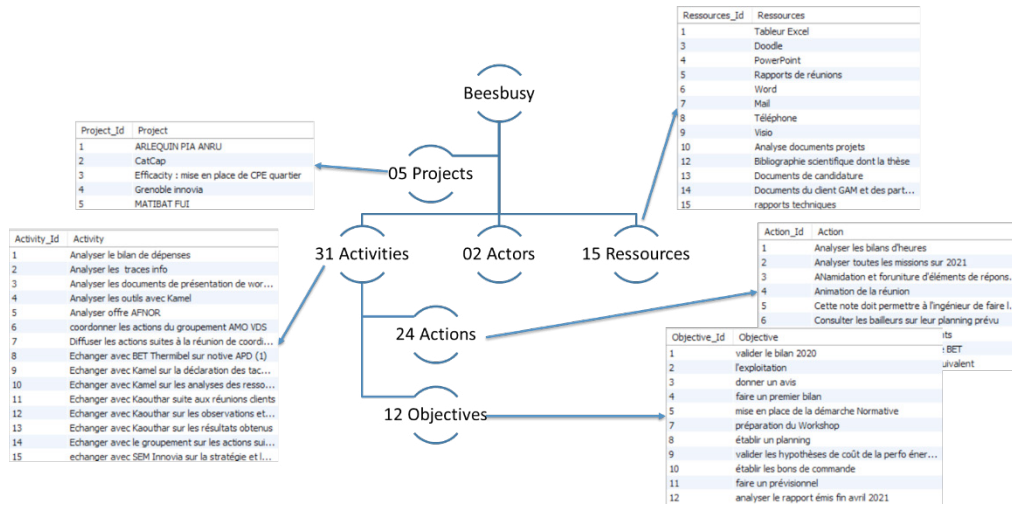


Figure 6. Beesbusy Extracted information

of 31 finalized activities as presented in Figure 6. The tables observed in Figure 6 have been generated as a result of the data processing step as defined in §4.2.

In terms of performance, the average of the calculated performance is 4.13. In this case, we can conclude that the actors were highly performing on the globality of the activities.

Figure. 7 is a capture of the global cartography constructed from the processed traces. This global cartography shows the degree of complexity and diversity of the data represented, as well as the difficulties of analysis from this overview.

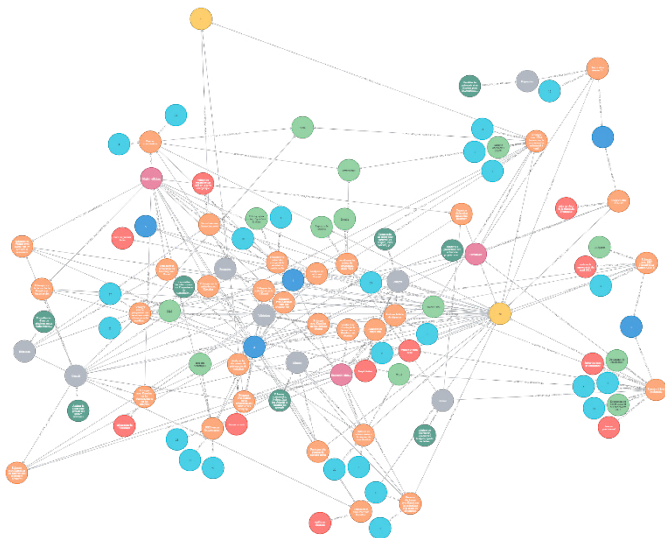


Figure 7. A global vision on the cartography

Once the data has been filtered through queries defined according to the needs of the analysis, we obtain more significant cartographies. For example, Figure 8 shows a view of the identified competence clusters. We can observe that across the 31 analyzed activities, only 03 competences clusters have been identified and are related the actor 01. These results have two possible explanations. First, information concerning the rest of the actor's activities are missing, and this is because the actor has not respected the given instructions for all the activities. Second, some activities are still in progress, and subsequently, no performance information can be calculated as mentioned in §4.3.

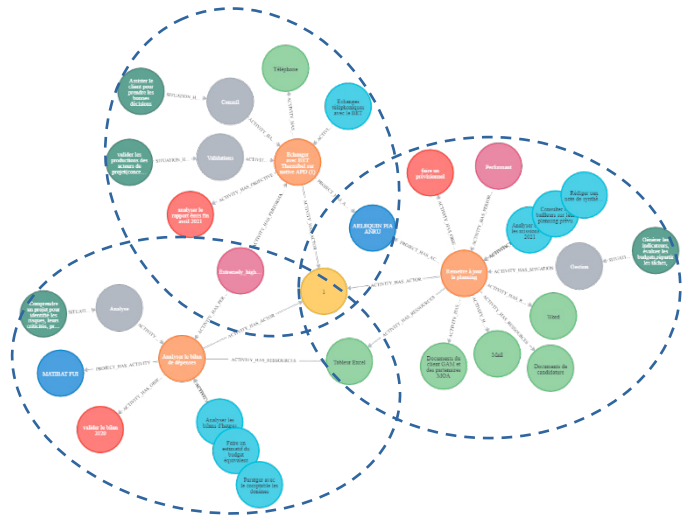


Figure 8. Identified competences clusters

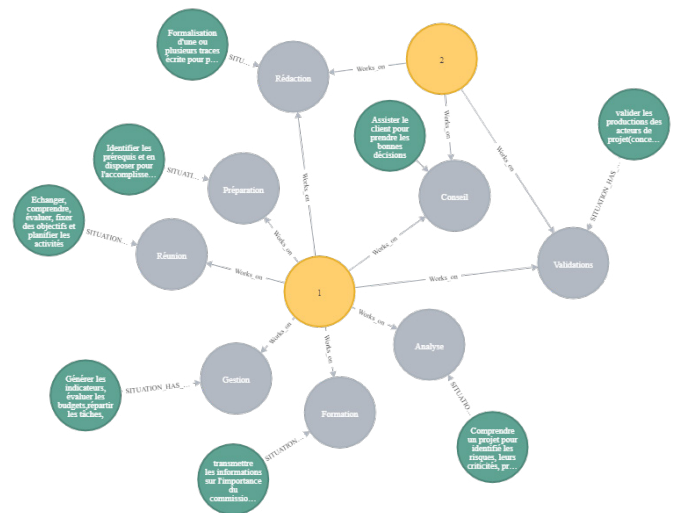


Figure 9. Actor's situations

In another view, the system permits the identification of the most frequent situation that each actor has been confronted with, as shown in Figure 9. This information can be used for example when building working teams to allow complementarity and experience sharing between actors.

## 6. CONCLUSIONS

We propose in this paper an activity-based system to identify and formalize the actual competences mobilized by the actors across daily activities. Our work contributes to the enrichment of the intangible capital of companies (competences). The application in industrial case study allows to validate our approach on identifying actual competences mobilized by actors. That said, there are unfortunately some weaknesses in our proposal. Out of all extracted data, only three competences have been extracted due to the lack of data for some activities. Besides that, weakness, if only one component is missing the competence cannot be built.

This paper introduces the basic concepts of our system and its overall architecture, in addition to a view on the resulting competence cartography. For the sake of space in this paper, the methods for passing through each part have not been covered in detail and will be the subject of a future paper.

Future work will focus on overcoming identified weaknesses and exploiting the competence cartography to manage the company's performance.

**Acknowledgment:** We would like to thank the ANR (french National Agency of Research) for having funding this work carried out in the framework of the CatCap project.

## REFERENCES

- Alecu, F. (2013) 'Tracking the Digital Traces', *Oeconomics of Knowledge*, 5(3), pp. 2–7.
- Belkadi, F., Bonjour, E. and Dulmet, M. (2007) 'Competency characterisation by means of work situation modelling', *Computers in Industry*, 58(2), pp. 164–178.
- Belmonte, J. L. and Segura-Robles, A. (2020) 'Machine learning and big data in the impact literature. A bibliometric review with scientific mapping in Web of science', *Symmetry*.
- Chouhan, V. S. and Srivastava, S. (2014) 'Understanding Competencies and Competency Modeling — A Literature Survey', *IOSR Journal of Business and Management*, 16(1), pp. 14–22.
- Coron, C. (2018) 'Big Data et pratiques de GRH', *Management & Datascience*.
- Courtin, C. and Talbot, S. (2005) 'An architecture to record traces in instrumented collaborative learning environments', *IADIS International Conference on Cognition and Exploratory Learning in Digital Age*.
- El Falaki, B., Idrissi, M. K. and Bennani, S. (2010) 'Formative assessment in e-learning: an approach to personalize diagnosis and adapt learning path', in *IADIS International Conference e-Society 2010*, pp. 391–395.
- Graczyk-Kucharska, M. et al. (2018) 'Model of competency management in the network of production enterprises in industry 4.0—Assumptions', *Lecture Notes in Mechanical Engineering*, (201519), pp. 195–204.
- Hand, M. and Gorea, M. (2018) 'Digital Traces in Context| Digital Traces and Personal Analytics: iTime, Self-Tracking, and the Temporalities of Practice', *International Journal of Communication*, 12(0), p. 17.
- Hecklau, F. et al. (2016) 'Holistic Approach for Human Resource Management in Industry 4.0', *Procedia CIRP*, 54, pp. 1–6.
- Kolibáčová, G. (2014) 'The Relationship Between Competency and Performance', *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis*, 62(6), pp. 1315–1327.
- Kou, Y., Jia, Z. and Wang, Y. (2013) 'A Comparative Research on Competency and Competence, Competency Model and Competence Model', in *Lecture Notes in Electrical Engineering*, pp. 681–693.
- Lund, K. and Mille, A. (2009) 'Traces, Traces d'interactions, Traces D'apprentissages Définitions, Modèles Informatiques, Structurations, Traitements et Usages', in (Eds.), J. C. M. & A. M. (ed.) *Analyse de traces et Personnalisation des EIAH*. Paris: Lavoisier-Hermes
- Miranda, S. et al. (2017) 'An ontology-based model for competence management', *Data & Knowledge Engineering*. Elsevier, 107, pp. 51–66.
- Nicolaescu, S. S. et al. (2020) 'Human capital evaluation in knowledge-based organizations based on big data analytics', *Future Generation Management*. Elsevier.
- Pinzone, M. et al. (2017) 'Jobs and skills in industry 4.0: An exploratory research', *IFIP Advances in Information and Communication Technology*, 513, pp. 282–288.
- Robert, G. (2012) 'A study on the perception of human resource executives on competency mapping for the superior results', *International Journal of Social Science Tomorrow*.
- Settoui, L. S. et al. (2009) 'A trace-based system for Technology-Enhanced Learning Systems personalisation', *Proceedings - 2009 9th IEEE International Conference on Advanced Learning Technologies, ICALT 2009*, pp. 93–97.
- SIMIC, M. and NEDELKO, Z. (2019) *Development of competence model for industry 4.0: A theoretical approach, Economic and Social Development*, (Esd 2019): 37Th International Scientific Conference on Economic and Social Development - Socio Economic Problems of Sustainable Development.
- Wang, N. (2016) 'Towards a Competency Recommender System from Collaborative Traces'. UNIVERSITY OF TECHNOLOGY OF COMPIÈGNE.
- White, R.W. (1959) 'Motivation reconsidered: the concept of competence', *Psychological Review*, 66(5), pp. 297–333
- Wilcox, Y. and King, J. A. (2013) 'A professional grounding and history of the development and formal use of evaluator competencies', *Canadian Journal of Program Evaluation*, 28(3), pp. 1–28.
- Zarka, R. et al. (2012) 'Trace replay with change propagation impact in client / server applications', in *IC 2011, 22èmes Journées francophones d'Ingénierie des Connaissances*, pp. 607–622.