



HAL
open science

Automation of competences extraction - HumaNeTrust: an experimental computer system used in an academic laboratory

Olivier Pouydebat, Christophe Courtin, Kamel-Eddin Bemmami, Elodie Gardet, Georges Habchi

► To cite this version:

Olivier Pouydebat, Christophe Courtin, Kamel-Eddin Bemmami, Elodie Gardet, Georges Habchi. Automation of competences extraction - HumaNeTrust: an experimental computer system used in an academic laboratory. CIGI-Qualita21: Conférence Internationale Génie Industriel QUALITA - Grenoble, May 2021, Grenoble, France. hal-03203592

HAL Id: hal-03203592

<https://hal.science/hal-03203592>

Submitted on 20 Apr 2021

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.

CIGI 2021

Automation of competences extraction

HumaNeTrust: an experimental computer system used in an academic laboratory

OLIVIER POUYDEBAT¹, CHRISTOPHE COURTIN¹, KAMEL-EDDIN BEMMAMI¹,
ELODIE GARDET², GEORGES HABCHI¹

¹ Univ. Savoie Mont Blanc, SYMME
5 Chemin de Bellevue - 74940 Annecy-le-Vieux
olivier.pouydebat@univ-smb.fr, christophe.courtin@univ-smb.fr,
kamel-eddin.bemmami@univ-smb.fr, georges.habchi@univ-smb.fr

² Univ. Savoie Mont Blanc, IREGÉ
4, Chemin de Bellevue - BP 80439 - 74944 Annecy-le-Vieux Cedex
elodie.gardet@univ-smb.fr

Abstract

In an organisation, competences form human capital, an intangible asset that is decisive for overall performance and strategic management. The aim of this paper is to understand, in an objective and undeclared way, how competences in an organisation evolve. Then this research proposes to fill this gap by setting up a computer system to provide an automated overview of competences (individual and collective). After defining the competences, we show that traces can be used to identify them. Based on the PROXYMA model, we switch to an actual dimension for the first time through the implementation of an application (HNT: HumaNeTrust) that allows to validate the different functions of a trace-based system. Then, the HNT application is experimented in an organisation, based on of the traces produced during publication of scientific articles. The results show that the graphs produced can be used by the organisation's managers to obtain an objective overview of the professional interpersonal relationships and areas of expertise held within the organisation (which are some of the facets of the complex concept of competence). We notice that the areas of expertise found with HumaNeTrust software are not exactly the same as those highlighted on the organisation's website. This research marks a step in the quest for an automatic competence mapping solution, to provide the structure's managers with an objective view of reality.

Keywords – Competences mapping, Collective and Individual competences, Human information, Knowledge management, Skills, Competencies, Digital interaction, Trace-based system, TBS, HAL.

1 INTRODUCTION

One of the major components of a company's value is the human capital that is a part of its immaterial capital [Boucher et al., 2007]. The strategic management of a company implies the management of this human capital that comprises the individuals' information. A clear and undistorted view of this information is crucial for the leaders of a company.

The use of software tools for strategic management could facilitate its extraction [Balog et al., 2012] especially in a turbulent environment forcing the different actors to adapt and develop their competences. For managers, the challenge is to be able to identify quickly and objectively the internal competences that can be mobilized for a project development. In addition, the competences mapping can help managers to reveal the competences that are missing or those that need to be strengthened (for example, through training or hiring). Beyond their usefulness for managers, individual and collective competences are essential to establish the strategic diagnosis of a company and thus establish the main orientations.

Despite the various definitions of competences, due to the polysemic nature of the term, most researchers agree that they are closely related to the activities [Aubret et al., 2005; Bernabé-Moreno et al., 2019; Gluga et al., 2011; Lingappa et

al., 2016; Ughetto and Paris-est, 2002] and include multiple facets in terms of knowledge, know-how and skills [Durand, 2015].

Like the mainstream [Kandampully, 2002], we consider that competences being human dependent and temporally situated imply trust level and regular update issues, respectively. The difficulty then lies in successfully operationalizing this vision of competences. One of the proposals often used, is to base the analyses on the competences declared by the individuals themselves. However, some researchers believe that a distortion can exist [De Siqueira and Gough, 2016], because there is a gap between declared and real competences.

To fill this gap, we propose to get an automated inventory of elements of competences by means of the publications activities' analyses. Most studies based on publications analysis are conducted to estimate the impact factor of publications [van Leeuwen and Wouters, 2017], or to specify parameters influencing this impact factor [Sahoo et al., 2018]. As our objective is concerned, it consists in getting this automated inventory through publishing activities' analyses mainly from the HAL (Hyper Articles on Line: <https://api.archives-ouvertes.fr>) platform. These analyses are carried out in a Trace-Based System (TBS) and involve an evaluation of the level of reliability of the used instrumentation technique [Courtin, 2008].

The two main concepts of this study are interactive digital traces and competences. This work is based on the experience acquired from previous research work on a TBS called BEATCORP [Chebil et al., 2011], where we observed the relevance of low level traces for identifying potential individual competences [Chebil et al., 2015]. In this study, we collect traces at the task level to underline the link between tasks and competences as presented later. Therefore, we have implemented a software, based on the PROXYMA model [Chebil et al., 2015], to collect and organize higher level traces in order to identify competences.

The underlying question is to know whether it is possible to extract individual or collective competences from a trace-based system. This work thus offers both academic and practical value to an automated vision of competences. First, the concepts associate the definition of competences with the main dimensions of characterization. Second, the notion of trace is explained and how it can be useful to track competences. Then on these theoretical bases, our project is described (the HumNeTrust software). The main objective is to collect, organize, identify and map competences from traces of activities. Thanks to the first acquired results, we present a first mapping of competences. After the results presentation, we conclude with an in-depth discussion.

2 COMPETENCE

A huge volume of research works were dealing with the individual competences until the last decade, producing an impressive number of competence definitions. Due to the polysemy of this term, there is no consensual and universally accepted definition [Aubret et al., 2005; Lingappa et al., 2016; Ughetto and Paris-est, 2002]. Before suggesting a definition, we present a picture of the main common features that emerge from this abundance of definitions.

The scientific terminology highlights different meanings for the following concepts: skill, competency and competence [Fazel-Zarandi and Fox, 2010; Le Deist and Winterton, 2005]. The people competencies designate the cause and effect relationship between the individual characteristics, observed during a high yield activity, and their successful and effective job performance. The people competences highlight the link between a complex set of knowledge, skills, motivation, attitude, etc. and their job fits within a specific work context. In short, "competence" describes what people can do, while "competency" focuses on how they do it [Le Deist and Winterton, 2005; Sanghi, 2016].

In accordance with the above-mentioned definition of competence, we identify three dimensions in the set of elements: "Knowledge", "Know how to do" and "Know how to be" [Durand, 2015; Le Deist and Winterton, 2005].

This paper deals with the concept of competence which involves the above-mentioned complex items to carry out specified actions in a specific context to achieve a S.M.A.R.T. (Specific, Measurable, Achievable, Relevant and Time bound) objective [Menon, Saraswathi and Karl, Judith and Wignaraja, 2009]. According to this definition, this section tries to specify the meaning of the following terms:

Knowledge: an information acquired during learning activities or/and a set of facts, principles, etc. acquired through human experience and thought. It may be explicit or tacit [Nonaka, 1995]. Individual's expertise fields are those areas in which the person has in-depth knowledge and for which the person is socially recognized (reputation, recommendation, publications)

as a reference person.

Skills: a specific aptitude for a task, sport, etc., the ability acquired through training and/or practical activities [Durand, 2015; Le Boterf, 1994; Le Deist and Winterton, 2005].

Attitudes: a personal set of features that lead to a particular behavior in various work situations [Durand, 2015].

Action: concrete manifestation of someone's activity [Durand, 2015].

Context: a set of elements related to human aspects, financial aspects, time, material resources, etc., which define a specific situation [Brownell, 2008].

The competences are revealed through action and recognized by the others. Many studies indicate that competences can only be approved if they are socially recognized. The action and the social recognition are thus both essential aspects to address the three dimensions of the competences. People competences emerge thanks to a good implementation of their own knowledge in the action [Durand, 2015; Le Boterf, 1994].

The number of skill classifications being high, we present below an excerpt reduced to the most common types found in the literature [Durand, 2015]:

- Individual competences: an individual's own skills. Individuals generally include these competences in his/her job description, his/her curriculum vitae or his/her LinkedIn© profile for example. The subtypes listed below are used commonly to distinguish individual's competences. Studies [Belkadi et al., 2007] strives to synthesize the publications on individual competences as following :
 - Hard skills: demonstrable skills attested by the diplomas, certificates, degrees, titles ...
 - Soft skills: knowing how to collaborate, how to adapt to a situation, how to negotiate ...
 - Mad skills: atypical skills, acquired through the practice of a sporting or creative activity, or an original passion ...
 - Collective competences: they are acquired by the quality of the interactions that take place between the individuals in a company. Collective competences are made up of two complementary parts, one referring to the individual and another one to the organization. Any modification of one of these parts modifies the collective competence itself [Boreham, 2004; Le Boterf, 2000; Macke et al., 2013]. Valérie Michaud [Michaux, 2005] provides a synthesis of the types of collective competences dynamics : i) Collective competences in the perspective of pooling, synergy effect. (project team, operational team), ii) Collective competences in the perspective of change (collective learning, adoption of a new technology, creation of new knowledge), iii) Collective competences in the perspective of sharing (free or guided) experiences. (knowledge sharing, reflection group, progress group, etc.), iv) Collective competences in the perspective of daily management of hazards: cooperation, transversal communication, inter-individual exchanges within a network.
- Organizational competences: a set of competences that allow performing actions whose objective is to

successfully accomplish the organized actions, but also to influence the way in which these actions are organized. These competences allow to organize the framework of action, and allow the evolution and adaptation of it to new constraints [Grant, 1991; Michaux, 2005].

Business owners, and more specifically in the human resource department, collect and organize this information according to an internal skills repository. Such repositories have inventory values because they contain the description of the positions and the jobs. These repositories become referentials in the companies and they are now the subject of many studies [Jamias and Oiry, 2013]. The ESCO survey accessible online provides a multilingual repository of individual skills in the form of an updated and effective taxonomy for the European labor market (ESCO is the multilingual classification of European Skills, Competences, Qualifications and Occupations: <https://ec.europa.eu/esco/>). The use of referentials is constantly increasing and a common frame of references with a shared representation can be defined by means of consensus or even negotiations. Such a representation would enable to identify both the activities to be carried out and thus the corresponding actions to do and the procedures to be applied to reach the objectives. This information can be input into software tools to assist the human resource managers during staff appraisal or the leaders to establish a strategic management.

As noted above, the holding of competences is evidenced through actions. Research and analysis of the traces of these actions will serve as evidence of possession of competences.

Competence having been defined, it now appears as a complex and composite concept, with a very varied typology. In light of this complexity, it seems clearer that managers in organizations need tools to help them identify the elements that make up the competences present in their structures. A competence extraction tool will be able to detect each of the facets that make up competences. It will provide a limpid, clear vision that is as exhaustive as possible.

3 TRACE

As it was mentioned in the previous section, competences may be identified through activities containing functional and business tasks. When executed in a Digital Working Environment (DWE), the corresponding actions leave digital interaction traces that can be analyzed. During the last two decades, there has been an increasing interest in analyzing activities especially in the learning management systems (LMS) in the form of observable objects. In the scientific literature, many research works deal with the notion of trace, but there is no standard definition to this day [Settouti et al., 2009]. As defined by [Settouti et al., 2009], traces are considered as a collection of temporally located observables which are any structured information stemming from the observation of human-machine interaction (Observable: structured information derived from observation of an interaction). These interaction traces are collected in a trace-based system (TBS) to be transformed into higher abstraction levels traces to identify the corresponding tasks listed in referentials. The functional and business tasks imply soft and hard competences related to behavior and application fields, respectively. These competences are listed in referentials similarly and may be thus identified through the actions of individuals or groups of users. Through the ESCO European referential, it is thus possible to identify the competences and

attitudes that are linked to the specific tasks. In conclusion, the interaction traces provide information on two of the three dimensions of competences: know how to do and know how to be. In terms of knowledge identification, [Barazzutti et al., 2015] presents an interactive tool, named "Transmut", to support knowledge discovery (KD). This tool enables the extraction of knowledge by means of traces processing following a cycle composed of three main phases. Firstly, traces are collected, selected, and transformed. Secondly, an algorithm is used to detect frequent sequential episodes considering the constraints specified by the analyst. Thirdly, the results of the operation are transmitted to the analyst who chooses the most relevant models according to his knowledge of the domain. In terms of competences and attitudes identification, there is no significant proposal of studies focusing on the exploitation of interactive traces. Nevertheless, some attitudes (soft competences) can be measured by the traces collected from different actions of user activity. H.Tadger's work on PBL (Problem Based Learning) environments for the improvement of students' soft competences on the basis of activity traces, shows that it is indeed possible to measure some soft competences (e.g. in communication, in time management, initiative, curiosity) [Tadger et al., 2020]. The system developed in H.Tadger's work under the name "SYDIS" (SYstem of Development and Improvement of Skills) allows to measure through traces (traces of direct actions; traces of interaction with the system) the level of these soft competences and their evolution, as demonstrated in the experimental study of the SYDIS.

As shown above, traces stemming from actions may be used to identify the three dimensions of competences. Through the ESCO standard, it will be possible to identify the competences and attitudes that are linked to the identified tasks. The basis of our approach lies on the exploitation of these traces.

The UML model presented figure 1 enables to describe interaction traces with various abstraction levels based on various granularity levels [Chebil et al., 2015]. A signal corresponds to an elementary action identified by a single event, such as save a page in a word processing software. A sequence has some duration and is composed of a chronological succession of signals and other sequences. These traces are stored in the database of a TBS based on the CARTE model (Collection, activity Analysis and Regulation based on Traces Enriched) [Courtin, 2008] and are exploited with several application programming interfaces (API) to be collected, transformed, queried, notified and displayed. The analysis of these traces needs the description of the context of the activities where these actions take place. These contextualized traces are organized in corpora containing context metadata and other resources of the corresponding activities, such as provided or produced documents. In this study, these corpora are managed with the BEATCORP system (Benchmarking platform for Analysis of Trace CORPora) based on the PROXYMA approach [Chebil et al., 2015] whose originality is to exploit traces regardless of their original format. Their description is not the subject of this paper, but the TBS presented below is based on the BEATCORP architecture.

Thus, with the information mined behind traces, we can use a competence model, to evaluate users' competences on some subjects. N.Wang [Wang et al., 2016] proposes "Action-Knowledge model" of competence permitting the evaluation of users competences (eg. Java programming; PHP; C#). This evaluation is based on an indicator of competence ($C\{i,j\}$)

denoting the competence of user j on subject i , calculated according to a statistical method used for the research of information in textual documents.

We can also use a competence model to extract company competences in a context of decision-aid to identify collaborative networks. A.Harb presented an approach of ontology enrichment for competence extraction implemented through the developed system "UNICOMP" [Harb et al., 2011] to identify collaborative networks. This approach focuses on extracting essential facets of competences within the company by analyzing traces of competences on its website.

The following section proposes a first experimentation based on the concept of traces and competences. Through this experimentation, we develop a trace-based system to identify the researchers competences in a laboratory.

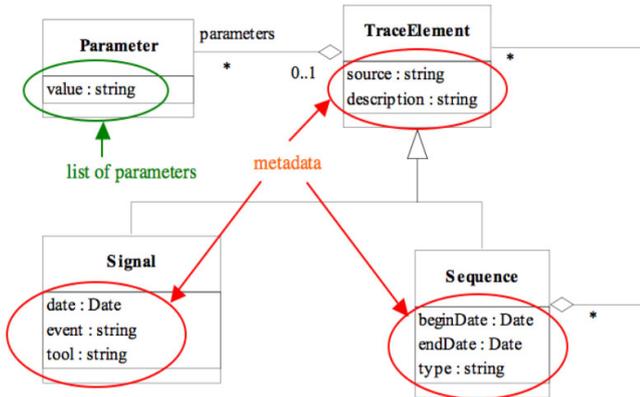


Figure 1 : Interaction Trace Model

4 FIRST EXPERIMENT : HUMANE TRUST

Our objective was dual: (i) to equip an organization with a competence mapping tool (ii) to implement a solution according to the concepts specified in the TBS BEATCORP. The HumaNeTrust (HNT) prototype was implemented in PHP and javascript with a MariaDB relational database. We chose to study the actions of publication of articles for the researchers of a laboratory. It was a very instructive phase because it allowed us identifying the effective difficulties, and especially the limits of this first proposal. This is what it will be detailed below.

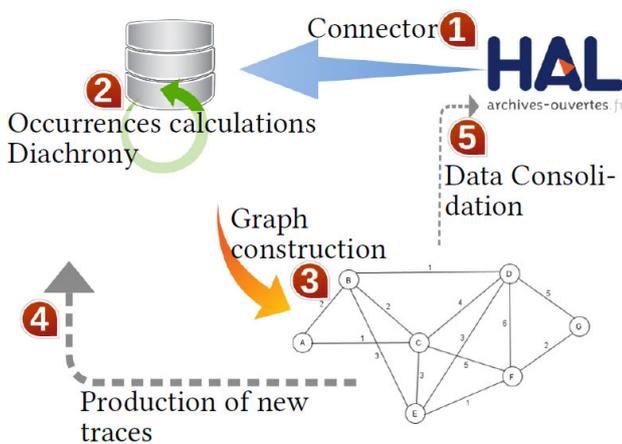


Figure 2 : HumaNeTrust process principle

The five main functions of HNT (Figure 2) are the following:

- Step 1: data recovery from HAL (a French repository

dedicated, among others, to research articles). A connector is in charge of retrieving the state of the data in HAL, at regular intervals, and storing this raw data in the local database.

- Step 2: the raw data are organized in such a way as to calculate the occurrences. (eg. Number of times a researcher has published using a given keyword. Number of times two researchers A and B have co-published.)
- Step 3: production of a graph illustrating these organized data. Users can interact with the graph by clicking on a node or edge. The user also has a search engine to find hidden terms in the graph.
- Step 4: the user can complete the displayed data with additional information that he can input. This information can be attached to a node or an edge of the graph.
- Step 5: this first experiment had the effect of highlighting some of the inconsistencies present in the HAL data source. The researchers themselves went to check the source and made corrections. It is a process of data consolidation.

The main functions of HNT being described; here are the details of the implementation.

4.1 Data acquisition

The HAL Restfull API (<https://api.archives-ouvertes.fr/>) is used to set up the functions that are specific to the connection to the data source (Mark N°1 in Figure 3). At this stage, data are stored as a time-stamped local buffer, in the original unmodified JSON format.

4.2 Data processing

The second program (HAL_DataCompilation.php) loads a raw JSON file, as retrieved from HAL and stored in the database by the first program. Then the program fills a set of relational database tables so that we can make powerful SQL queries (Mark N°2 in Figure 3). Three tables store organized source data: the first for all authors, the second for all publications and the third for all keywords. The calculated occurrences are stored in three specific tables: the first contains the co-publications occurrences, the second contains keywords occurrences used by authors, and the third contains the occurrences of co-use between two keywords.

4.3 Graphs processing

In a third program, data sets are built with these pre-calculated data to be displayed in the form of graphs. (Mark N°3 in Figure 3). There are two types of graphs in JSON format:

Graphs whose nodes are authors: a node represents a person. The diameter of the node carries the information of the number of times the author has co-published. The edges of this graph indicate co-publication between two authors. The thickness of the edge is proportional to the number of co-publications between two authors.

Graphs whose nodes are keywords: a node represents a keyword. The diameter of the node carries information on the number of times the keyword has been used in the laboratory's publications. The edges of this graph indicate the use of two keywords in co-publication. The thickness of the edge is proportional to the number of times a keyword has been used at the same time as another keyword in a publication. The graphs can contain a large number of nodes and edges that can degrade the fluidity of the consultation interface, the

readability of the entities and the general ergonomy of the tool. Some functionalities enable thus to compute graphs with all the nodes representing data stemming from HAL queries or only a part of them.

4.4 Data viewing

Several graph sets are stored in JSON format and the users may load one of them on their Web browser by a forth program (Mark N°4 in Figure 3). They can explore the interactive displayed graph by selecting nodes or edges to obtain their properties. A search engine allows the user to look for an information in the graph, and the resulting graphic entity can be emphasized by a zoom functionality. We use a library to spatialize the nodes according to the occurrences carried by the edges. Higher is the occurrence, higher is the attraction force between two nodes. Initially, we choose arbitrary to distribute evenly the nodes on a circle. Then, the spatialization program calculates the relative positions of the nodes according to the values of occurrences carried by the edges between these nodes. There are various methods of spatialization [Cogis and Schwartz, 2018; Thomas, 2015], and their limits are shown in the table below:

Table 1: Algorithm of spatialization.

Algorithm	Year	Limit
Fruchterman Reingold	1991	maximum 1000 nodes
Yifan Hu	2005	maximum 100000 nodes
Force Atlas 2	2010	maximum 1000000 nodes

The SIGMA-JS (<http://sigmaj.s.org/>) library is compatible with a graph in JSON input format, and enables to implement Force Atlas 2 in javascript on the client side.

4.5 Inputting Data in Networks

In a future development, we will complete the graph interface by functionalities allowing collecting user input, associated with nodes and edges. The objective is to collect three types of new data:

Feedback: This type of input will allow us to create new data, or categorize existing data based on nodes and edges. We will be able to build input forms specific to an information gathering campaign, and record the new data in our database. (Mark N°7 in Figure 3).

Consolidation: When users see an error (eg. in the spelling of their name), they can immediately notify it or even correct it themselves. Providing them with the tools to perform these operations could well be possible from HNT. (Mark N°6 in Figure 3).

Recommendation: It is also planned to offer to the HNT users some tools allowing them to evaluate (eg. from 1 to 5 stars) a level on a graphical entity (node or edge). For example, their level (self-evaluated) of expertise on a specific domain. In addition, they could evaluate the elements of their N+1 level network, if the graph represents their social network, centered on themselves (ego-centric network). (Mark N°5 in Figure 3).

4.6 Overall vision

The solution combines four functionalities: i) Data acquisition, ii) Data processing, iii) Graphs processing, iv) Data watching. Figure 3 shows these different levels of. It is planned to add a JSON data management module to HNT in order to deal with the buffering of raw data. We investigated the Elasticsearch stack (Elastic, Beats, Logstash: <https://www.elastic.co/>) solution which would be appropriate due to:

Its scalability: the possibility to easily increase the storage capacity.

The tools availability and compatibility: in particular Beats and Logstash are very good solutions to fill the Elasticsearch base. It will be able to load a local buffer from various and heterogeneous data sources.

Finally, arrows start from the visualization tool (HNT) or from the future recommendation system (Mark N°5 in Figure 3) and go until the HNT base (Mark N°7 in Figure 3) and the HAL base (Mark N°6 in Figure 3). These arrows represent flows that do not yet exist in an operational way, and whose purpose will be to: i) Consolidate HAL data from inconsistencies highlighted in graphs, ii) Produce new data from the interactions carried out in the graphs (data entry on nodes or edges, suggestion, etc.).

Since July 2019, two graph consultation terminals are in place in the laboratory and can be used by researchers to explore keywords used in publications. For ethical and confidentiality reasons, author's graphs are not displayed publicly. This technical platform will be used as a basis for the first experiments aiming to automate the extraction of competences from traces present in HAL. The graphs produced will be able to be used by the laboratory management to know in an objective way the interpersonal professional relationships and the competences held within the organization.

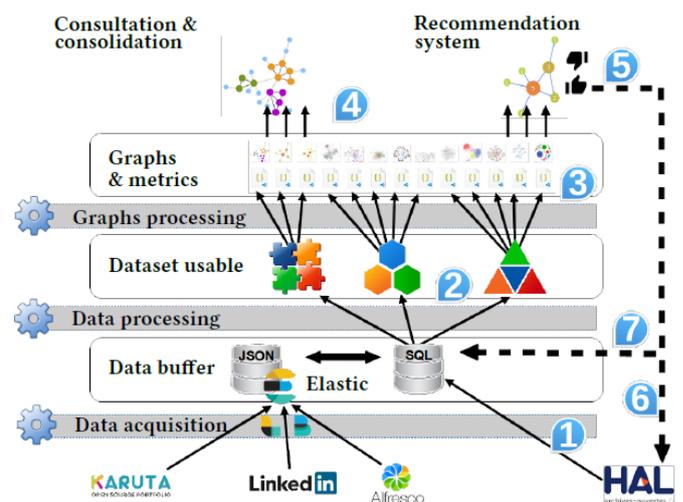


Figure 3: The HumaNeTrust architecture

5 RESULTS: HUMANETRUST SOFTWARE

Finally, the HumaNeTrust software combines the two main concepts presented in sections 2 and 3 of this work: competences and traces. The results of these concepts are present in HumaNeTrust: figure 4 shows the software, and the architecture has been described in section 4. The used graphs are based on an export of data on 2020/02/14, processing 1209 publications. Table 2 shows the graphs volumetry.

Table 2 : Graph volumetry at 2020/02/14.

	Total publications	Total nodes	Total Edges
keywords graph	1209	1914	9029
authors graph	1209	914	972

This implementation is structured according to the Proxyma models.

This application is composed of the four zones:

The main background screen: The graph is displayed with an adjusted zoom factor. The node size changes according to the

value to be represented (number of publications for an author, or number of occurrences for a keyword). The links between nodes (edges) are of constant thickness. However, we can associate to this thickness either the number of co-publications between two authors, or the number of co-uses of two keywords in the publications. It is possible to click on two types of entities in the graph: i) Click on a node: highlights the subgraph of the nearest direct neighbors. This action also updates the contextual zone (left side), ii) Clicking outside the graph, in the background: cancels the current selection and returns to the total selection of the graph. This action also updates the contextual area (on the left side).

A search tool: It is possible to enter directly a term (a keyword or an author's name) and if the item is present in the graph, it will be automatically centered and selected.

A contextual zone: three subsections are present

- Graph caption: not finalized, this section will display the graph metadata.
- Node information: not finalized, this section will display the properties of the clicked node. Later, it will also be possible to enter data to be associated with the currently selected node.
- Neighbors of keyword: this section displays a cloud of keywords. Terms present in the cloud are the keywords of the neighbors connected to the currently selected node.

Navigation tools: global view, zoom in and zoom out. These tools complement the actions available directly in the graph, such as zooming in and out with the mouse wheel, and pan.

5.1 Uses

Two large touchscreens have been installed on two of the laboratory's sites: "point A" and "point B". In both cases, these spaces are strategic places in the laboratory where many researchers pass by, and can take a break to exchange ideas. The produced graphs answer the following questions that researchers ask themselves when discovering the software: What are the most relevant keywords? To which keywords are they related? Which authors use them in the laboratory?

At first glance, researchers are surprised to see the keyword "Simulation" being the most used in the publications. This term largely dominates the others. However, this field of expertise is not highlighted by the laboratory in its communication (website, social networks, etc.). It is nevertheless the field of expertise of the laboratory most shared by researchers and the most cited in their publications.

These three questions are the main ones that generate discussion. Unfortunately, for reasons of respect for individual rights, the graph for answering the third question is not available on public touch screens. However, during demonstrations on private workstations, we were able to observe researchers using the authors' graph. Their first reaction is to look for themselves in the graph. Some authors analyze their personal network, and mentally reconstruct the reasons for the different links. Finally, and this was a desired goal from the beginning of the experiment, some researchers get that they have multiple identifiers in the graph. The beneficial effect of these findings is that the researchers correct the data source (Mark N°5 in Figure 3).

It is planned to anonymize the graph of the authors and they will be represented with a unique identifier. The user can then click on this identifier and send a message to the author. This message states that a user has wished to identify him, and ask

him in return to reveal his identity. Answers to other questions have been requested. For example, which authors publish articles with the most popular keywords? Which authors introduce new keywords?

This last question introduces an important planned evolution: the management of the chronology of appearance of keywords in publications. For this reason, we will have to compute new graphs according to the following sequence:

1. Sort publications by increasing publication date.
2. Identify the emergence of new keywords.
3. Identify the authors of these publications.

The keywords graph can then be coupled to a time slider, which will allow visualizing the growth of the graph progressively as the new keywords appear.

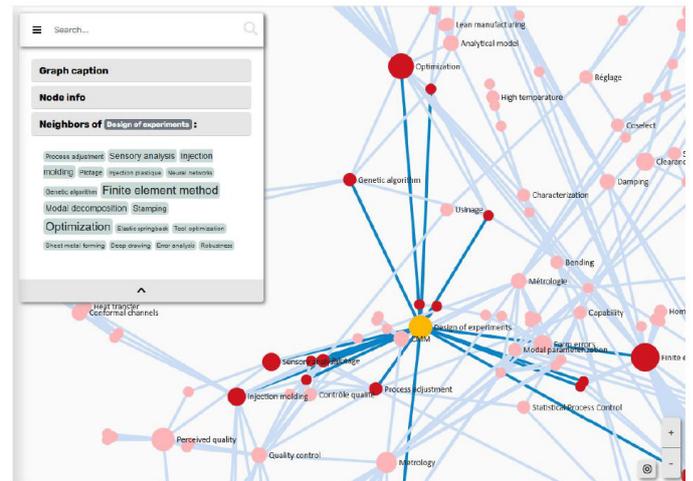


Figure 4: HumaNeTrust user interface. Graph of keywords used in the laboratory's publications, and their co-use links

5.2 Current limitations and prospects

Beyond the storage and processing mechanisms presented, it is important to keep in mind that the goal is to be able to extract a competences mapping from the system.

A critical analysis of the existing system is thus necessary:

Area of expertise and competence: what is the difference between publishing in an area of expertise and having recognized competence in the subject? As we saw in section 2, the field of expertise is one of the facets of competence, included in knowledge, one of the three dimensions that makes up competence (knowledge, expertise, interpersonal skills). Our solution for extracting competences identifies some of the information relating to the structure, but does not offer an exhaustive view of the individual and collective competences actually present. However, this information is valuable because it allows us to deduce all the activities of the researchers, complementary to the publication work around these fields of expertise: teaching, search for partnerships, participation in professional networks, bibliographic research, modelling, English writing, collaborative work, etc. For example, some of these activities address indirectly the expertise facets of competences that are linked to the writing of scientific articles. There is therefore additional work to be done to enrich the competence profile of individuals with other sources of information that it will be possible to cross-reference. On the other hand, it will certainly be useful to calculate and quantify the level of expertise. It would certainly be relevant to compare

these areas of expertise with a notoriety calculation (i.e. H index: <https://arxiv.org/abs/physics/0508025>) on these keywords in order to add the "social recognition" facet of competence.

Genericity: the built system is highly oriented on a type of data (scientific publications or studies). However, it will be easy to reconfigure the system to consider data from an e-mail box, LinkedIn, etc.

Diachony: build graphs provide a static view of the status of the publications. The basic functions are implemented for the dynamic construction of a network on user actions. The literature shows the established interest of the process-oriented competence vision [Bayad et al., 2006]. Chronological graphs will be completed with metrics calculations [Dakiche et al., 2019], and differential vision between two dates.

Going further in graph use: the analysis of the graphs was not carried out as thoroughly as desired. It would be very useful to calculate metrics based on graph theory and more particularly on the theory of strong links and weak links [Saeedi et al., 2018]. Highlighting structural holes and structural bridges, calculating "clustering coefficients" and "connection coefficients", are so many indicators whose relevance is not in doubt. The navigation in the graph, i.e. the process of exploring the graph, proposed by the application is quite basic. A development has been initiated to propose a billiard type navigation [Ventocilla et al., 2017].

Missing references: lot of noise is generated by the missing references.

- People repository: the use of an LDAP directory, coupled with text analysis, shows that duplicates were reduced to more than 98%. This experimental development has not been completed, and has not been integrated into the HNT source code.
- Keywords repository: the back up of HNT with an ontology system [Ilahi et al., 2014] would make it possible to simplify the graphs by making thematic groupings.

A problem of scale and granularity: there are many other competences in a laboratory than those reflected in the publications of the researchers. These competences are not proved, neither in action nor socially. Yet these two elements are inseparable from the definition of competence, as presented in section 2. A suitable system should enable to: i) put actions back at the center of the analyses, ii) increase the variety of observed actions, iii) recover a form of social validation relative to the quality of the implementation of actions.

Through this self-criticism, HumaNeTrust appeared to be an interesting step towards the use of traces to manage the competences of the laboratory. This step was necessary to partially validate the feasibility of the implementation of the PROXYMA models [Chebil et al., 2015].

As we have just underlined in this section, these first results make us want to go further and perfect our base to add functionalities to it. The initial reactions are encouraging, and show us that the needs of structure managers, in terms of competences mapping, are considerable.

6 CONCLUSION AND PERSPECTIVES

In this document, we validate an approach based on traceability to identify automatically the human capital of organizations and partially individual and collective

competences. Based on the academic work that led to the Proxima model, the HumaNeTrust application was set up and implemented in two public places. Initial feedback indicates that HumaNeTrust offers a new vision of an organization's human capital: competences appear in an objective and factual manner. The work of data collection and organization is well automated, and snapshots can be taken at regular intervals in order to analyze the processes of building or losing competencies.

We have moderated the interpretation of our results, which we certainly consider encouraging, but which we would like to supplement with a number of improvements, which we propose to you in the following descriptions. It is interesting to note that the areas of expertise found with the HumaNeTrust software are not exactly the same as those highlighted on the organization's website. This is a form of evidence that there is a gap between the competences known and displayed by an organization and the competences actually held by its human capital.

We also note that the HumaNeTrust software provides us with an objective view of the interpersonal professional relationships present in the organization's activities. This display opens up ethical questions that we will have to answer with adapted technical solutions.

The work of defining competences proposed in this article shows that knowledge and expertise (extracted from the fields of expertise) are only part of competences. These elements are already rich in information. In the proposed approach, the tracing of actions as well as social recognition are lacking to prove that an individual has acquired competences. Further research will have to be carried out to identify the sources of real competences present in a group of individuals, capable of relying on a wider variety of actions.

This first experience has allowed us to measure the limits of our approach regarding the studied data and the implemented functions in the TBS. These limitations allow us to identify some research questions and the associated technical problems. One of these issues concerns the link between the concepts of trace and competence, both related to activities. In other words, can competence be considered as a high level of abstraction of the trace or can a trace be considered as evidence of competence? Secondly, how can these traces be better used for intelligent competence management in organizations?

7 REFERENCES

- Aubret, J., Gilbert, P., and Pigeyre, F. (2005). *Management des compétences: réalisations, concepts, analyses*. Dunod.
- Balog, K., Fang, Y., De Rijke, M., Serdyukov, P., and Si, L. (2012). Expertise retrieval. *Foundations and Trends in Information Retrieval*, 6(2–3), 127–256.
- Barazzutti, P., Cordier, A., Fuchs, B., Barazzutti, P., Cordier, A., Fuchs, B., and Tool, I. (2015). *Transmute : an Interactive Tool for Assisting Knowledge Discovery in Interaction Traces*.
- Bayad, M., Boughattas, Y., and Schmitt, C. (2006). Le métier de l'entrepreneur : le processus d'acquisition de compétences. *8ème Congrès International Francophone En Entrepreneuriat et PME (CIFPME)*, 16.
- Belkadi, F., Bonjour, E., and Dulmet, M. (2007). Competency characterisation by means of work situation modelling. *Computers in Industry*, 58(2), 164–178.
- Bernabé-Moreno, J., Tejada-Lorente, Á., Herce-Zelaya, J.,

- Porcel, C., and Herrera-Viedma, E. (2019). An automatic skills standardization method based on subject expert knowledge extraction and semantic matching. *Procedia Computer Science*, 162, 857–864.
- Boreham, N. (2004). A theory of collective competence: Challenging the neo-liberal individualisation of performance at work. *British Journal of Educational Studies*, 52(1), 5–17.
- Boucher, X., Bonjour, E., and Grabot, B. (2007). Formalisation and use of competencies for industrial performance optimisation: A survey. *Computers in Industry*, 58(2), 98–117.
- Brownell, J. (2008). Leading on land and sea: Competencies and context. *International Journal of Hospitality Management*, 27(2), 137–150.
- Chebil, H., Courtin, C., and Girardot, J.-J. (2015). A case study to validate the PROXYMA approach to share and analyse contextualised interaction trace corpora in a TEL environment. *International Journal of Learning Technology*, 10(4), p.291-312.
- Chebil, H., Courtin, C., and Girardot, J.-J. (2011). BEATCORP: une plate-forme de benchmarking pour l'analyse de corpus de traces. *Conférence Environnements Informatiques Pour l'Apprentissage Humain (EIAH)*, 10.
- Cogis, O., and Schwartz, C. (2018). *Théorie des graphes* (Cassini (ed.); Poche). Collection L.
- Courtin, C. (2008). *CARTE : AN OBSERVATION STATION TO REGULATE ACTIVITY IN A LEARNING CONTEXT*. *†. Celda*, 191–197.
- Dakiche, N., Benbouzid-Si Tayeb, F., Slimani, Y., and Benatchba, K. (2019). Tracking community evolution in social networks: A survey. *Information Processing and Management*, 56(3), 1084–1102.
- De Siqueira, J. R., and Gough, M. J. (2016). Correlation between experience targets and competence for general surgery certification. *British Journal of Surgery*, 103(7), 921–927.
- Durand, T. (2015). L'alchimie de la compétence. *Revue Française de Gestion*, 41(253), 267–295.
- Fazel-Zarandi, M., and Fox, M. S. (2010). Reasoning about skills and competencies. *IFIP Advances in Information and Communication Technology*, 336 AICT, 372–379.
- Gluga, R., Kay, J., Lever, T., and Lister, R. (2011). *An architecture for systematic tracking of skill and competence level progression in Computer Science*.
- Grant, R. M. (1991). The resource-based theory of competitive advantage: implications for strategy formulation. *California Management Review*, 33(3), 114–135.
- Harb, A., Hajlaoui, K., and Boucher, X. (2011). Competence mining for collaborative virtual enterprise. *IFIP Advances in Information and Communication Technology*, 362 AICT, 351–358.
- Ilahi, M., Cheniti-Belcadi, L., and Braham, R. (2014). Formal competence-based assessment: On closing the gap between academia and industry. *ACM International Conference Proceeding Series*, 581–587.
- Jarnias, S., and Oiry, E. (2013). Vers un repérage des types de référentiels de compétences. *@Grh*, 8(3), 11.
- Kandampully, J. (2002). Innovation as the core competency of a service organisation: The role of technology, knowledge and networks. *European Journal of Innovation Management*, 5(1), 18–26.
- Le Boterf, G. (1994). *De la compétence : essai sur un attracteur étrange*. Editions d'Organisation.
- Le Boterf, G. (2000). *Construire les compétences individuelles et collectives*. Ed. d'Organisation.
- Le Deist, F. D., and Winterton, J. (2005). What Is Competence? *Human Resource Development International*, 8(1), 27–46.
- Lingappa, D. M. S., De Fransisci Morales, G., and Gionis, A. (2016). Extracting skill endorsements from personal communication data. *International Conference on Information and Knowledge Management, Proceedings, 24-28-Octo*, 1961–1964.
- Macke, J., Sarate, J. A. R., and Vallejos, R. V. (2013). Collective Competence and Social Capital Analysis in Collaborative Networks. *Journal of Systemics, Cybernetics and Informatics*, 8(3), 18–23.
- Menon, Saraswathi and Karl, Judith and Wignaraja, K. (2009). Handbook on planning, monitoring and evaluating for development results. *UNDP Evaluation Office, New York, NY*, 68(3), 10.
- Michaux, V. (2005). Performance collective et compétences individuelle, collective et organisationnelle : construction d'une grille d'analyse unifiée. *Agrh*, 1–26.
- Nonaka, I. (1995). *Knowledge and Management: Chapter 2 of Ikujiro Nonaka and Hirotaka Takeuchi, The Knowledge Creating Company: How Japanese Companies Create the Dynamics of Innovation*. New York: Oxford University Press.
- Saedi, A., Peukert, E., and Rahm, E. (2018). Using Link Features for Entity Clustering in Knowledge Graphs. In *Springer International Publishing AG, part of Springer Nature 2018* (Vol. 1, Issue March). Springer International Publishing.
- Sahoo, J., Mohanty, B., and Dash, I. (2018). An analytical study on the publication pattern and impact of top research papers: A case study of information processing and management. *Library Philosophy and Practice*, 2018(October), 1–22.
- Sanghi, S. (2016). *The handbook of competency mapping: understanding, designing and implementing competency models in organizations* (P. Enclave (ed.); Second Ed.). SAGE publications India.
- Settoui, L. S., Prie, Y., Marty, J.-C. C., Mille, A., Prié, Y., Marty, J.-C. C., and Mille, A. (2009). A trace-based system for Technology-Enhanced Learning Systems personalisation. *Proceedings - 2009 9th IEEE International Conference on Advanced Learning Technologies, ICALT 2009*, 93–97.
- Tadger, H., Lafifi, Y., Seridi-bouchelaghem, H., and Gülseçen, S. (2020). Improving soft skills based on students' traces in problem-based learning environments. *Interactive Learning Environments*, 0(0), 1–18.
- Thomas, S. A. (2015). *Data Visualization with JavaScript* (No Starch).
- Ughetto, P., and Paris-est, U. (2002). *Service Competence: an Overview*.
- van Leeuwen, T. N., and Wouters, P. F. (2017). Analysis of Publications on Journal Impact Factor Over Time. *Frontiers in Research Metrics and Analytics*, 2(April).
- Ventocilla, E., Bae, J., Riveiro, M., and Said, A. (2017). A billiard metaphor for exploring complex graphs. *CEUR Workshop Proceedings*, 1798, 37–40.
- Wang, N., Abel, M., Barthès, J., Negre, E., Wang, N., Abel, M., Barthès, J., Negre, E., Wang, N., Abel, M., Barthès, J., and Negre, E. (2016). *Mining user competency from semantic trace*. May 2015.