XXXIX CONFERENZA ITALIANA DI SCIENZE REGIONALI

SOFT SKILLS OF THE FUTURE: A HETEROGENEOUS APPROACH BASED ON DATA MINING TECHNIQUES

Elena Coli[1], Silvia Fareri[2], Gualtiero Fantoni[3], Filippo Chiarello[4], Giacomo Petrini[5]

ABSTRACT

The advent of Industry 4.0 paradigm is shifting the attention more and more towards what the future of work will be. Researchers are currently working on which will be the strategic skills to succeed in the labour market and the professional profiles more required by companies. The "macro-world" of skills 4.0 could be divided in two classes: Hard Skills and Soft Skills.

We define 4.0 Hard Skills as "a set of technical skills that are necessary to create, use and maintain the enabling technologies of Industry 4.0" and Soft Skills as "a set of transversal skills, which could be divided on personal, interpersonal and technical skills".

The proposal work focuses on the detection of Soft Skills 4.0, starting from different databases of profiles and competences (e.g. ESCO, O*NET, Scientific Literature...).

Moreover, the final objective of the research is to analyse soft competencies in relation to Industry 4.0 context, finding out which of these are likely to become fundamental in the new industrial paradigm.

The analysis starts with the collection of Soft Skills from a series of heterogeneous sources, in order to consider the phenomenon from different points of view, thus including soft skills belonging to different contexts.

Once a consistent list of homogeneous soft skills would be obtained, the most important Soft Skills in the Industry 4.0 context are ranked.

However, each skill could become obsolete when a machine is able to perform it better, faster or cheaply. To understand the risk of obsolescence a series of queries on scientific paper literature have been performed and their degree of automation measured.

In this way it would be possible to highlight the skills that are still resilient to automation. Finally, a proposal *list of Skills of the Future (*the ones with the major impact on Industry 4.0) have been defined.

[1] Fondazione Giacomo Brodolini, Largo Lucio Lazzarino, 56122, Pisa , elenacoli@hotmail.it

[2] UniMore, Fondazione Marco Biagi, Fondazione Giacomo Brodolini, Largo Marco Biagi, 10, 41121 Modena, silvia.fareri@gmail.com

[3] Dipartimento di Ingegneria Civile ed Industriale, Largo Lucio Lazzarino, 56122, Pisa, g.fantoni@unipi.it

[4] Dipartimento di Ingegneria dell'Energia, dei Sistemi, del Territorio, delle Costruzioni, Largo Lucio Lazzarino, 56122, Pisa filippochiarello.90@gmail.com

[5] Consorzio QuiNN, Piazza Francesco Carrara, 19, 56126 Pisa, g.petrini@consorzioquinn.it

1. Introduction

Industry 4.0 is defined as a trend of automation and data exchange based on the use of new technologies and their interconnections (Bernard M.). Because of its innovative power, the phenomenon has been compared with the other three revolutions in industry: the steam power to mechanize production, the electric power to create mass production, the electronics and robots to automate production. Therefore, Industry 4.0 or the digital revolution is frequently called “the Fourth Industrial Revolution”.

Moreover, while the scientific interest in technological aspects of Industry 4.0 is constantly growing, the understanding of the future works and professional roles is slowed down by the heterogeneity, complexity and static nature of job description systems. These issues are usually addressed by qualitative methods thus making the results uncertain and incomplete.

In Italy and in general throughout the world, the concept of Industry 4.0 is getting more and more attention. The theme of competences is of fundamental importance for the community, as the competences represent fundamental drivers of the change that the labour market is facing. This growing interest implies the need to predict which competences and which professional profiles, seen as a set of competences, will be crucial in the future.

On one hand, Industry 4.0 implementation may have negative impact on specific professional figures, which could be substituted by automated systems (robots, chatbot, Artificial Intelligence software). On the other hand, it will be possible to assist to a significant change of the key competences needed to face the phenomena: our research mostly focused on this second prospect.

Most of all, the role of soft skills is acquiring more and more strategic relevance and they are going to make the real difference in the digital era.

The paper is structured as follows: firstly, a literature review helps the reader to contextualize the work and its contribution to the scientific community. Secondly, the methodology adopted to reach the analysis objective is deeply described through a workflow. In the end, some examples of extractions and visualization are shown, in advance to results discussion and future developments formulation.

2. Literature Review

2.1 Industry 4.0 Phenomenon

Rapid changes in working environments and newly adopted technologies have a profound impact on economic development (Markusen A., 2004). While technological progress has generally augmented productivity, it has also caused structural unemployment, inequality and social imbalance (MacCrory F. *et al.*). In such uncertain economic environment, standardizing data about the skills and attributes of occupations is a valuable asset for firms, managers and employers to access basic information about the requirements and needs of the job market (Hilton M.L., 2010).
Moreover, the epochal phenomenon called “Industry 4.0” is radically changing the existing social and economic structures that characterize the global market (Last C., 2017). However, its impact is still uncertain and not punctually predictable.
The differences between this new production paradigm and the previous ones are the exponential growth, the global scope and the certain (and still not clearly defined) impact on social, economic and political structures.
The expression "Industry 4.0" has been used for the first time during the "Hannover Messe", a fair on the industrial technologies, in 2011. One year later, the Working Group on Industry 4.0 presented a set of Industry 4.0 implementation recommendations to the German federal government.
The advent of Industry 4.0 strongly influences the Italian productive context. Small and medium-sized enterprises (SMEs), or companies with less than 250 employees, represent 99,9% of the Italian companies and 80% of employment.
The major problem of SMEs is finding resources for innovation and, in a context of continuous evolution and change, this problem paralyzes companies, preventing them from fully participating in technological evolution.
That is the reason why, the demand for skills 4.0 often remains too weak and limited to the needs of large companies. The rest of the Italian economy is concentrated in traditional sectors with low productivity, where there is little demand for high-level skills, with about 85% of Italian companies that are small and mainly family-owned.
Italy is therefore in a state of equilibrium, characterized by supply and demand of skills tending towards levelling down, in a vicious circle that has negative repercussions on growth and use of new technologies. In order to face this problem, the Italian Government, at the beginning of 2017, emitted a governmental plan for the development of Industry 4.0, that consists in two phases (Ministero dello sviluppo economico, 2017).
The first phase provides important tax relief for companies that would have invest in innovation and new machinery. This phase led, in the past year, to a relaunch of investments.
The second phase has started at the beginning of 2018: managers and workers will soon be able to re-train themselves on manufacturing 4.0, thanks to the tax incentives on training.
In particular, the plan for the second phase provides four pivotal points:

1. spreading the Industry 4.0 culture through "Digital School" and work-school alternation;
2. developing Industry 4.0 competences through university courses and dedicated higher technical institutes;
3. funding the Industry 4.0 research by strengthening PhDs;
4. creating "Competence Centre" and "Digital Innovation Hub".

Integrating the new paradigm in already existing organizations (small, medium or large) and raising awareness toward workers are two important topics to address for a smooth and successful implementation of the new paradigm. Developing skills and competencies is one of

the most effective way for the actual development, diffusion and implementation of Industry 4.0 in our society.

2.2 The "Skill" concept in years: a wide overview

A wide overview of the concept of "skill" in years is necessary for performing an effective research analysis. Moreover, it is fundamental to provide a clear definition of what the term "skill" really means. In order to do that, it is useful to define the terms "knowledge", "ability "and "attitude" too. Table 1 shows different definitions and their correspondent references.

Table 1: Definitions of Attitude, Knowledge, Ability and Skill or Competence

	SKILL, ATTITUDE, KNOWLEDGE, ABILITY	
	Definition	Given by
Attitude *(Know how to be)*	A learned predisposition to respond in a consistently favorable or unfavorable manner with respect to a given object.	Fishbein and Ajzen, 1975
Knowledge *(Know)*	Knowledge indicates the result of assimilation of information through learning. Knowledge is the set of facts, principles, theories and practices related to a field of study or work; knowledge is described as theoretical and/or practical.	Recommendation of the European Parliament and Council of 22 April 2008 - European Qualifications and Qualifications Framework
Ability *(Know-how)*	Ability indicates being able to apply knowledge and use know-how to complete tasks and to solve problems; abilities are described as cognitive (use of logical, intuitive and creative thinking) and practices (which imply manual skills and the use of methods, materials, tools).	
Skill or Competence	Skill indicates the ability to use *knowledge, abilities* and *personal, social and/or methodological skills*, in work or study situations and in professional and/or personal development; the skills are described in terms of responsibility and autonomy.	

Source: authors' elaboration

Typically, in literature and in the common practice (jobs atlases, job descriptions, etc.), skills are organized in two macros groups: the hard skills and the soft skills. Their definitions are provided on Table 2.

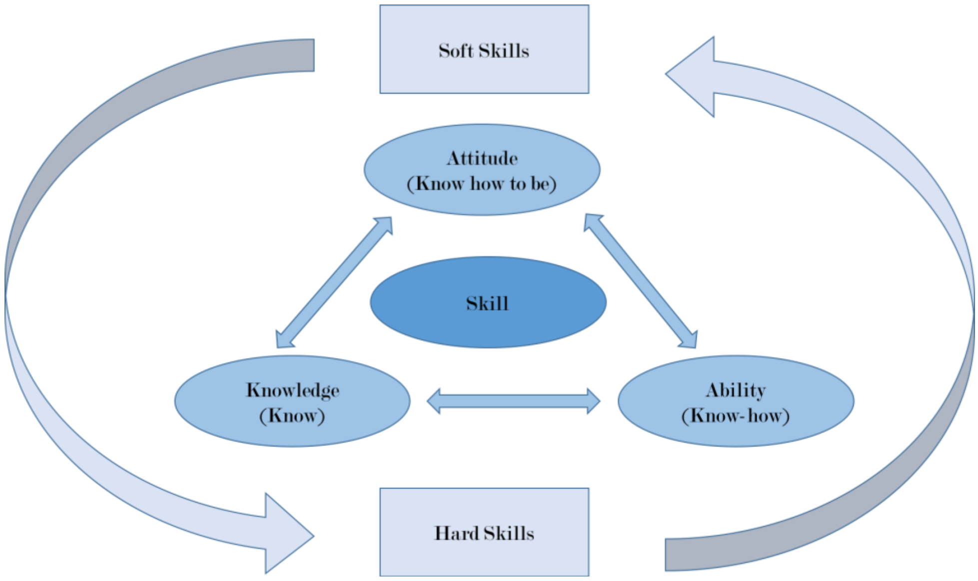
Table 2: Definitions of Hard Skills and Soft Skills

	HARD AND SOFT SKILLS	
	Definition	Given by
Hard Skills	Hard Skills make reference to *technical-professional knowledge*, tightly correlated to the nature of the job and the developed activities.	QuInn, 2017
Soft Skills	Soft Skills represent a dynamic combination of *cognitive skills*, *interpersonal skills*, *intellectual skill* and *practical skills* alongside *ethical values*. They enable individuals to adapt and behave positively in order to effectively face the challenges of daily and professional life.	Haselberger et al., 2010
	Soft Skills are people skills supported by their emotional intelligence that helps them to behave in a socially acceptable way and to adapt to a social environment, so that others feel comfortable in their company and vice versa.	Verma, 2009
	Soft Skills are those "transversal" skills not related to a specific sector or job, which favor the growth of companies and the employability of people; therefore they represent an added value for the competitiveness of companies.	Dell'Amico, 2016
	Soft Skills are a set of non-technical skills and knowledge that support effective participation at work. They are not specific to the type of work and are strongly related to the qualities and personal attitudes, social skills and management. Because of their intangibility, some of these capabilities are difficult to quantify, recognize, evaluate and develop.	European Centre for the Development of Vocational Training (Cedefop)
	Soft Skills are interpersonal skills, which is the ability to establish and maintain relationships with other people.	PMI, 2013
	Soft Skills are personal traits, goals, motivations and preferences that are considered important in the labor market but also at school and in other areas. They are predictive of success in life.	Heckman & Kautz, 2012
	Soft Skills are complementary to the technical-professional skills developed during the academic and/or work experience and are useful for knowing how to adapt to the professional sphere and the changes that characterize it. They are personal resources necessary to manage and deal with the different aspects of their work, such as social relationships in group and relationship situations, possible stressful moments or the need to be creative or reactive in the face of problems or unforeseen events.	Pezzoli et al., 2017

Source: authors' elaboration

Finally, Figure 1 graphically describes the relationship between the aforementioned definitions.

Figure 1: Graphical representation of the definitions



Source: authors' representation

Based on the enabling technologies defined by Boston Consulting Group, it is clear that the advent of Industry 4.0 pushes the development of both hard and soft skills to face the epochal labour challenges. Moreover, dealing with soft skills is even more complicated, because of their relationship with personality traits.

2.3 Skills in Industry 4.0 Environment

The radical technological change caused by the Fourth Industrial Revolution will affect the whole industrial environment and will radically transform, in several different ways, the world we live in.

Consequently, the revolution is not just referred to the availability of new sophisticated machines, but also to the deep reconsideration of workers roles.

On one hand, Industry 4.0 implementation may have negative impact on specific professional figures, which could be substituted by automated systems (robots, chatbot, Artificial Intelligence software). In support of this theory, Osborne and Frey implemented a methodology to estimate the probability of job computerisation using a Gaussian classifier (Frey B. C., Osborne M., 2016) . The results are quite pessimist: they distinguished between high, medium and low risk of computerisation, and, according to their estimates, around the 47% of jobs is in the high risk category. On the other hand, another possible impact is represented by a significant increase of worker competences and the emersion of new professional profiles.

Also World Economic Forum estimated that Industry 4.0 will delete 5 million job places (WEF, 2016). In contrast to this point of view, there are many careers that could be generated by digitalization: the workers who could be replaced by robots are a small percentage, while most of them will have to change their skills in order to adapt themselves to a new way of working, in order to be able to use technologies and to interpret data and results.

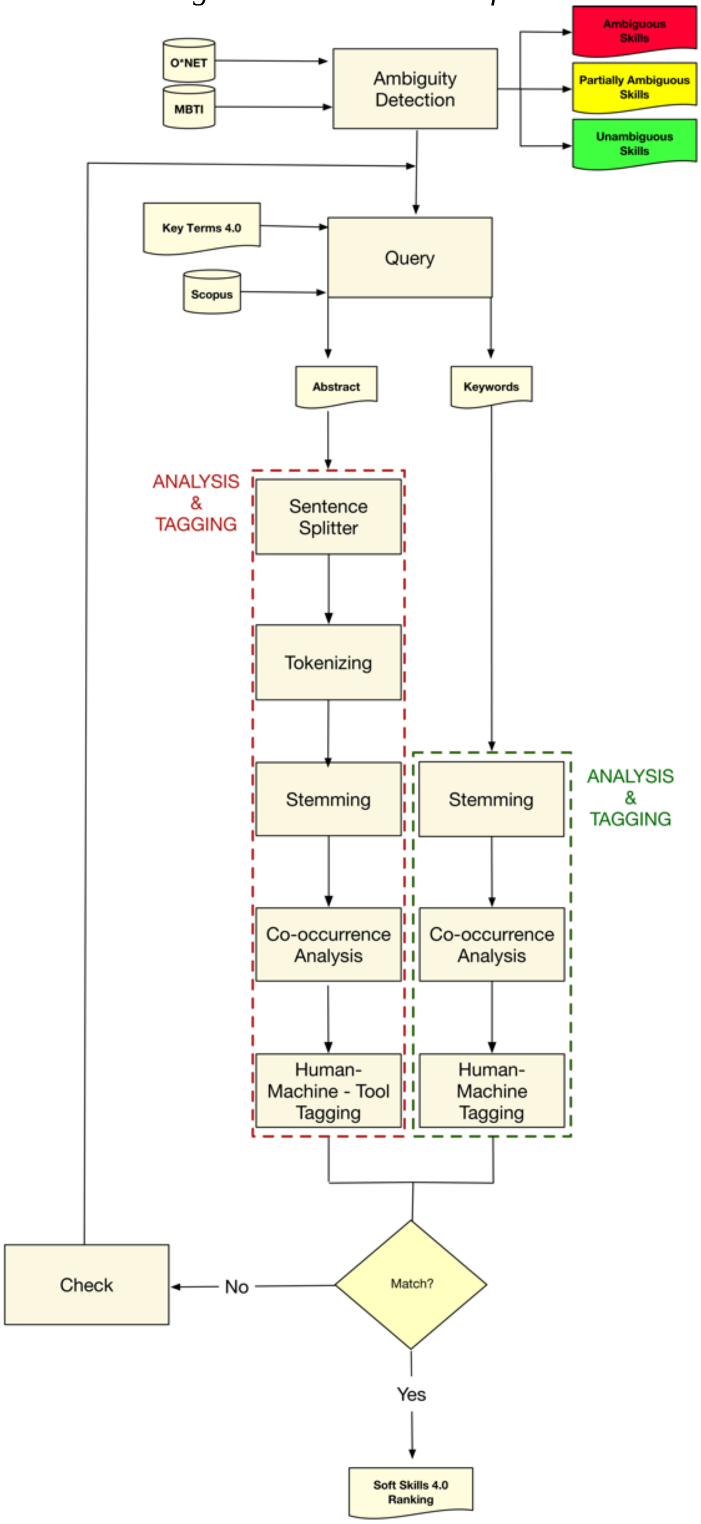
What we expect is not only a gradual substitution of man by machines but also an integration of new skills to existing professional profiles and the creation of completely new profiles. In parallel, a likely disruption of a large number of jobs. (Rainie L., 2017)

Therefore, the technologies involved in Industry 4.0 impact in a remarkable way on the human capital, demanding suitable competences to properly manage them.

The existing connections among skills has been used to explain several dynamics: the workers' transition between occupations, the way cities acquire comparative advantage in new skills, and how individual occupations change their skill requirements (Alabdulkareem, A, 2018). Our work focuses on understanding which skills are the ones most related to 4.0 topic.

3. Methodology

Figure 2: Research Workflow



Source: authors' representation

The Industry 4.0 paradigm could be considered a recent phenomenon. Thus, forecasting its impact without using a large amounts of data is a quite difficult task.
Moreover, the use of dated sources could show a retrospective vision rather than providing a forward-looking information. For these reasons, it was decided to gather the skills from different sources, such as scientific publications concerning Industry 4.0 enabling technologies. The true purpose of using publications was capturing current trends of scientific research.

3.1 Data extraction and ambiguity detection

First of all, a comprehensive analysis of ISFOL, ESCO and O*NET databases was performed. After that, O*NET database was selected, because of its widely diffusion within industries and its key role assumed on the frequently mentioned work about the future of employment by Osborne and Frey.
Moreover, the skills contained in O*NET are all expressed in English, a concrete advantage for a subject - Industry 4.0 - which finds a lot of space in the international literature. It also has a univocal correspondence in ISFOL[6] database, which contains information on Italian workers about the same competences.
Among the information contained in O*NET, the fields that are most commonly used by previous studies are those which describe the skills, abilities and typical work activities corresponding to each occupation (W.L. Woon *et al*, 2015).
Thus, the analysis was conducted on a subset of O*NET database (https://www.onetonline.org), consisting of "Cognitive Abilities" (i.e. abilities that influence the acquisition and application of knowledge in problem solving), "Basic Skills" (i.e. developed capacities that facilitate learning or the more rapid acquisition of knowledge), "Social Skills" (i.e. developed capacities used to work with people to achieve goals), "Complex Problem Solving Skills" (i.e. developed capacities used to solve novel, ill-defined problems in complex, real-world settings), "Systems Skills" (i.e. developed capacities used to understand, monitor, and improve socio-technical systems), "Resource Management Skills" (i.e. developed capacities used to allocate resources efficiently) and "Work Styles" (i.e. personal characteristics that could affect how well someone performs a job).
Then, an expansion of the list was carried out by using an additional source: the official descriptive document of the Sixteen Personality Types, defined through the Myers-Briggs Type Indicator (MBTI) (I. Myers and P. Myers, 2013).
The duplicate values were deleted and the expressions were compared, in order to eliminate those having the same meaning.
Only the most appropriate expression for each skill was maintained. The selection criteria adopted for the aforementioned skimming phase are reported in order of precedence. In particular, the skills selected were: (1) those belonging to the O*NET list, (2) those made by only one term and (3) those most frequently found.
Once a complete list of soft skills was identified, a Soft Skills 4.0 ranking was built, quantifying the correlation between the identified skills and the Industry 4.0 topic. Scopus database was used as data source to achieve the previous goal. The analysis faced several critical issues, mainly due to the complex formulation of several skills, often subject to polysemy and therefore not reliable for the creation of the queries.
For this reason, the skills were divided into three groups, according to their length and polysemy: a first group consisting of ambiguous and strictly polysemic competences, of variable length; a second group of not polysemic skills, with a long (more than three words)

[6] Istituto per lo sviluppo della formazione professionale dei lavoratori

and complex formulation; a third group consisting of not polysemic competences with a length between one and three words.
After that, some of the skills were reprocessed in order to avoid their ambiguity, mainly using synonyms for not polysemic ones and stemming (reduction to the root) for polysemic ones. On one hand, the use of synonyms and stemming increased the recall. On the other hand, the next step in narrowing the field of papers, increased accuracy, taking into consideration the results related to the topic only.

3.2 Query creation and occurrences collection

Moreover, the query on Scopus was always built applying the same pattern[7].
The first critical issue to be faced was the poor occurrence of several queries, which did not reach the minimum acceptable threshold (set at 100 in relation to occurrences distribution drop) to assure the significance of the following automatic analysis. For this reason, they were not analysed. The second problem identified was the meaning ambiguity of several skills found in scientific publications, which were not uniquely referable to human beings. Therefore, it was necessary to develop a methodology to contextualize the analysis results. The relation between the skill under consideration and both human competence and machine feature, were quantitatively evaluated, in order to attribute a “human” or “machine” label to each skill.

3.3 Abstracts and keywords analysis and tagging

The first context analysis was conducted by extracting the Keywords (Index and Author) associated with the different papers; the two types of Keywords were voluntarily kept separate to check whether they provide the same information. The Keywords were manually labelled as “human” and as “machine”. The weight with which the keywords associated with “human” (”machine”) influenced on each skill was finally checked. When the two categories of Keywords indicated the same label (H or M), the process stopped; in case of discrepancies, the analysis was repeated on a finer level of detail, by extracting and analysing the Abstracts. In fact, they represent a larger source of text and, consequently, of information too. The text of each abstract was divided into single words or expressions of 2 or 3 words, and manually labelled H-M again. Finally, the human component associated with each skill, called Human Corrective Factor (HCF), was calculated as a percentage. Whenever the HCF value exceeded the value (μ - σ) of the distribution, the skills were labelled as HUMAN, conversely whenever the HCF value was lower than the value (μ - σ), the skills were labelled as MACHINE. In case of intermediate values, it was not possible to exactly label the skill.

3.4 Final ranking creation

The final Soft Skills 4.0 Ranking was built after the calculation of a third reliability index (ABS)[8]. This index quantifies the reliability of the occurrence value. High occurrences and high HCF correspond to high reliability values of label attribution, while lower occurrences and lower HCF correspond to medium or low reliability values of label attribution.

---

[7] i.e. inserting the competence (reworked or not), followed by the term “skill” and its synonyms (stemmed to include also their plurals), and by a set of ten “digital words”: five synonyms of the concept of the Fourth Industrial Revolution and five technologies closely related to the concept of Industry 4.0. These words were obtained through a process of expansion and extraction of terms and technologies typical and centred on the topic Industry 4.0 (Chiarello F. *et.al*, 2017).
[8] Occurrence * |0,5-HCF|

4. Results

4.1 Data extraction and ambiguity detection

The Soft Skills were automatically extracted from the aforementioned O*NET database categories.
Then, an expansion of the list was carried out by extracting soft skills from the descriptions of Sixteen Personality Types, defined by using the Myers-Briggs Type Indicator.
The Myers-Briggs Type Indicator (MBTI) and the Jung Type Indicator (JTI) (C. Jung, 2017) are two personality indicators. Their purpose is identifying several psychological characteristics, through appropriate psychometric questionnaires. The Myers-Briggs approach was chosen because it has been used for a long time and therefore it is easier to find in literature. The sixteen profiles textual descriptions were analysed and the soft skills contained within them were extracted.
Then, the duplicates were deleted and the two lists were compared, in order to find expressions with the same meaning. The last step was choosing the best term or expression for each skill having more than one way to be expressed.
After this process, a final list of soft skills was obtained.
In order to quantify the correlation between the identified skills and the Industry 4.0 topic, it was necessary to reformulate some skills that presented a certain level of ambiguity (possibility of a double interpretation) in their formulation. The risk is carrying out an analysis by using a too general expression. An example is the skill "Writing", which is not polysemic, but could be used in very different contexts. The first step was therefore identifying the ambiguous skills. In order to identify the different levels of ambiguity, a traffic light logic was used. The skills were highlighted in three colours, according to their linguistic criticalities.
The disambiguation phase was particularly complex, in fact a third of the skills were characterized by polysemy.
The Table 3 shows an extract of the Traffic Light Logic Tagging.

Table 3: Extract of the Traffic Light Logic Tagging

TRAFFIC LIGHT LOGIC TAGGING
Abstraction
Active Learning
Active Listening
Adaptability
Address Complexity
Altruism
Ambiguity Tolerance
Analytical Thinking
Attention to Detail

Source: authors' elaboration

While for the green skills the formulation was kept unchanged, a reformulation was necessary for the red and orange skills, in order to make the next stages more efficient.

4.2 Query creation and occurrences collection

Then, the queries for Scopus search were defined.
For each skill, the number of identified papers (the occurrence) was used as score 4.0.
The Table 4 shows an extract of the result.

Table 4: Extract of Soft Skills Occurrence

ORIGINAL SKILL	SCOPUS OCCURRENCE
Instructing	1885
Planning	1757
Decision Making	1520
Innovation	1281
Visualization	1259
Reliability	1145
Flexibility	1101
Networking	795
Programming	597
Precision	588
Problem Solving	548

Source: authors' elaboration

The high ranking positions are occupied by skills which concern the ability to training other people or to planning.
The results related to skills as “problem solving” and “decision making”, often evoked by the literature, are very interesting topics. In fact, they play a key role for the change that arises from the interaction between new technologies and emerging professional profiles.
The "problem solving” and the "decision making" could be found in the first positions of the ranking. This is clearly not surprising, considering their complementarity in dynamic and challenging working environments. For this reason, the worker 4.0 should have not only solid technical and scientific skills and knowledge, but he should promptly manage unexpected events at once, basically using intuition, flexibility and self-confidence.
Finally, it is possible to find the true managerial skills in the middle of the ranking. It could be said that the change will probably cause an indefinite phenomenon of skill disruption, but according to our result we could assume that it will not affect managerial roles.
Aware of the limits given by the linguistic processing of these skills, it is possible to make some considerations. The cognitive skill “visualization”, almost univocally attributable to machines, produced high occurrences. The central role of data analysis in the digital age is frequently mentioned in literature. For this reason, it is essential to be able to take advantage of ad hoc visualization techniques (i.e. graphs, histograms, Word Clouds, etc.). In fact, Data Visualization, the heart of Business Analytics, assumed the role of a real language: the best way for visualizing the analysis outputs certainly represents a strategic choice. Indeed, the interpretation of the results could change radically in relation to their visualization methods.
Thus, it can be said that the occurrence of each skill is partly due to the presence of papers indicating the link between the skill and the human, and partly to the presence of papers indicating the link between the skill and the machine. To conclude, there was a significant risk that the occurrence on Scopus referred to machine skills and not to human skills.

For this reason, a first level context analysis on papers' Author and Index Keywords was carried out.

4.3 Abstracts and keywords analysis and tagging

The skills were divided into two categories, in relation both to the occurrence drops and to the absolute consistency of occurrences number:

- occurrence greater than 100: the data were thick enough to allow an effective semantic analysis. In order to evaluate the accuracy of the occurrences, a first level context analysis on Keywords was performed (the method will be clearly explained in the next paragraph);
- occurrence less than 100: the method did not provide significant results, for this reason the skills were considered not analysable.

Then, the Author and Index Keywords were quantified, by counting the "n" occurrences of the same keyword in the different papers, and tagged manually, as terms that can be associated with human (H) or machine (M).
The two Keywords were voluntarily kept separated, in order to verify if they communicated the same information.
An extract of Author Keyword H-M tagging for "Project Management " is shown in Table 5.

Table 5: Extract of Author Keywords H-M Tagging for "Project Management"

	df_stem	n	tag
1	autom	14	M
2	big data	10	M
3	cloud comput	9	M
4	simul	5	M
5	knowledge manag	4	H
6	optim	4	H
7	product	4	H
8	systems integr	4	M
9	cad	3	M
10	engineering manag	3	H

Source: authors' elaboration

The Human-Machine percentages were then calculated in relation to the total number of keywords associated with each skill.
In the case the two Keywords provided conflicting information, it was necessary to perform a second level context analysis on papers' Abstracts. The Abstracts were tokenized and the words most related to the skill were selected and evaluated H-M manually.
A third variable was added to the Human and Machine ones, Tool, which actually represents the bridge between the two previous entities, capturing, in some cases, the tools that allow humans to manage the machines. The Table 6 shows an extract of Abstracts H-M tagging for "Project Management".

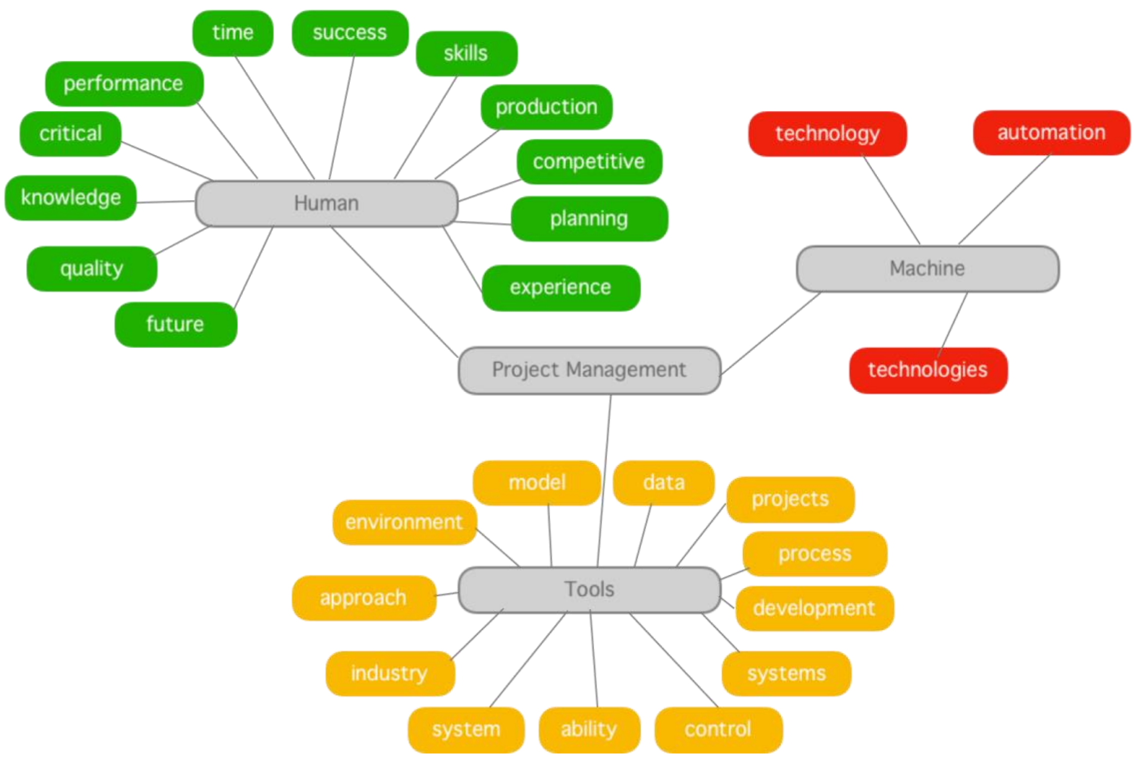
Table 6: Extract of Abstracts H-M Tagging for "Project Management"

	project_management	tag
process	0,99	T
development	0,99	T
projects	0,99	T
automation	0,989	M
ability	0,989	T
control	0,989	T
technology	0,988	M
systems	0,988	T
competitive	0,988	H
time	0,988	H
skills	0,987	H
technologies	0,987	M
industry	0,987	T

Source: authors' elaboration

The data were then displayed, allowing to make different considerations. First of all, the number of terms around a variable represented its relevance, allowing the evaluation and consequent selection of the most reliable between the two keywords percentages. Secondly, the cloud of concepts that revolves around tools allowed to visualize how the Human and Machine entities communicate with each other.
An example of graphic representation for the skill "Project Management" is shown in Figure 3.

Figure 3: Graphic Representation for "Project Management"



Source: authors' representation

Finally, the Human Corrective Factor (HCF) was calculated as a percentage and the final label (H-M) of each skill was identified.

4.4 Final ranking creation

The final Soft Skills 4.0 Ranking was built after the calculation of the reliability index (ABS). The intervals of ABS index (high, medium and low) were chosen based on results performance: the two thresholds were set in correspondence of two quantitative distribution drops.The Table 7 shows the 4.0 Soft Skills Results.

Table 7: 4.0 Soft Skills Results: Final 4.0 Ranking, HCF, ABS and Reliability Level

ORIGINAL SKILL	ORIGINAL OCCURRENCE	HCF	TAG	ABS	RELIABILITY
Visualization	1259	16%	M	422	HIGH
Monitoring	1677	26%	M	410	HIGH
Decision Making	1520	30%	M	308	HIGH
Instructing	1885	34%	H-M	304	HIGH
Networking	795	14%	M	290	HIGH
Planning	1757	34%	H-M	279	HIGH
Flexibility	1101	31%	M	205	HIGH
Reliability	1145	32%	M	201	HIGH
Precision	588	20%	M	177	HIGH
Innovation	1281	43%	H-M	90	MEDIUM
Respect	446	30%	M	88	MEDIUM
Initiative	395	28%	M	85	MEDIUM
Expertise	365	27%	M	84	MEDIUM
Reasoning	390	30%	M	77	MEDIUM
Systems Analysis	429	36%	H-M	61	MEDIUM
Abstraction	235	25%	M	59	MEDIUM
Popularity	227	25%	M	57	MEDIUM
Project Management	397	64%	H-M	56	MEDIUM
Cooperation	421	38%	H-M	51	MEDIUM
Leadership	218	72%	H	47	MEDIUM
Responsibility	359	39%	H-M	39	MEDIUM
Originality	276	36%	H-M	39	MEDIUM
Effort	198	34%	H-M	32	MEDIUM
Confidentiality	108	21%	M	31	MEDIUM
Adaptability	172	33%	M	30	MEDIUM
Resilience	129	28%	M	28	MEDIUM
Creativity	128	71%	H	27	MEDIUM
Concentration	183	40%	H-M	19	LOW
Just in Time	121	64%	H-M	17	LOW
Problem Solving	548	53%	H-M	16	LOW
Team-Working	109	64%	H-M	16	LOW
Judgment	124	39%	H-M	14	LOW
Management of Personnel Resources	139	57%	H-M	10	LOW
Mathematics	208	53%	H-M	6	LOW
Negotiation	166	53%	H-M	5	LOW
Multitasking	116	52%	H-M	2	LOW

Source: authors' elaboration

Analysing the results, the skills associated with decision making, monitoring, visualization and networking are strongly related to topic 4.0 and have a preponderant machine component.
The results about "decision making", apparently counterintuitive, may have a precise reason: in papers, machine learning used as support tool in automatic decision making is a topic whose attention is growing really fast.
At the same time, the results in terms of instructing, planning and innovation show that the adoption of the Industry 4.0 production paradigm needs extremely high and complex skills, which require high training. Problem solving 4.0 and judgement 4.0 underline the importance of making decisions quickly, based on collected data and on their often non-linear projections.
Although there is extensive room for improving the present methodology, it can be stated that a resilient profile 4.0 is a figure that is basically interdisciplinary, with strong and synergistic soft skills (ranging from the resolution of complex problems to marked critical thinking), necessary to manage and control "the Ecosystem 4.0".
However, the analysis results should be interpreted very carefully, because they can be influenced by the difficulties faced during the work. In fact, only about 40 skills of the starting data set were considered analysable after the occurrence exploration on Scopus.

5. Conclusion and Future Studies

The present work could be seen as the first step of a long journey. As stated in the introduction, professional profiles 4.0 need is increasingly emerging in the digital economy: on one side, there is the necessity of having new multiple figures able to manage Industry 4.0 phenomena, on the other side there is the need to update the skills of existing workers.
For this second aspect, thanks to the data-mining exercise, it was possible to extract focal competences for Industry 4.0. For several starting skills, the exploration reported a number of evidence too low to make any reliable deduction, so a further effort will be spent to refine the analysis. In fact, the work will require further efforts and, perhaps, the use of additional sources such as patents, curricula and job offers.
As a future development, one of the most interesting aspects is trying to automate and standardize the Human-Machine tagging procedure. In this sense, future works will certainly be focused on improving and enriching the methodology performed.
After that, we hope to identify the resilient and obsolescent job profiles using a combination of several tools (such as Soft Skills list and Technimeter © (Chiarello F. *et al*, 2017)) projected on textual profiles descriptions from additional sources, such as ESCO and the "World Robotics Industrial Robots and Service Robots" (https://ifr.org/worldrobotics).
In conclusion, the research goal was defining a quantitative approach for identifying Soft Skills of the future. The data-driven approach and the achieved results open the way for a future methodology extension (and improvement), through the integration with new data sources.

6. References

Marr, Bernard. *Why Everyone Must Get Ready For The 4th Industrial Revolution.*

A. Markusen (2004). *Targeting occupations in regional and community economic development.* J. Am. Plan. Assoc. 70 (3) 253–268.

F. MacCrory, G. Westerman, Y. Alhammadi, E. Brynjolfsson. *Racing with and against the machine: Changes in occupational skill composition in an era of rapid technological advance.*

M.L. Hilton, N.T. Tippins (2010). *A Database for a Changing Economy: Review of the Occupational Information Network (O*NET)*, National Academies Press.

Last, C. (2017). *Global Commons in the Global Brain.* Technological Forecasting and Social Change, 114, 48-64.

Ministero dello Sviluppo Economico (2017). *Piano nazionale Impresa 4.0.*

Carl Benedikt Frey, Michael A. Osborne (2016). *The future of employment: How susceptible are jobs to computerisation?*

World Economic Forum (2016). *The Future of Jobs.*

Rainie, L., & Anderson, J. (2017). *The Future of Jobs and Jobs Training.* Pew Research Center, (May), 1–95.

Alabdulkareem, A., Frank, M. R., Sun, L., Alshebli, B., Hidalgo, C., & Rahwan, I. (2018). *Unpacking the polarization of workplace skills.* Science Advances,4(7).

W.L. Woon, Z. Aung, W. AlKhader, D. Svetinovic, M.A. Omar (2015). *Changes in occupational skills-a case study using non-negative matrix factorization*, in: Neural Information Processing, Springer pp. 627–634.

I. Myers and P. Myers, Gifts differing (2010). *Understanding personality type.*

Chiarello, F., Trivelli, L., Bonaccorsi, A., Fantoni, G. (2017). *Development of an enriched dictionary for semantic analysis of documents in the Industry 4.0 domain.*

C. G. Jung, (2014). *Psychological types.*

Chiarello, F., Trivelli, L., Bonaccorsi, A., & Fantoni, G. (2018). *Extracting and mapping industry 4.0 technologies using wikipedia*. Computers in Industry, 100, 244-257.