Big data and skills frameworks to determine recommended profiled of soft skills for IS development

Vera Pospelova
Universidad de Alcalá
Alcalá de Henares, Spain   *vera.pospelova@uah.es*

Inés López-Baldominos
Universidad de Alcalá
Alcalá de Henares, Spain   *ines.lopezb@edu.uah.es*

Luis Fernández-Sanz
Universidad de Alcalá
Alcalá de Henares, Spain   *luis.fernandez.sanz@uah.es*

Ana Castillo Martínez
Universidad de Alcalá
Alcalá de Henares, Spain   *ana.castillo@uah.es*

Abstract

Information systems development (ISD) is a social activity where non-cognitive skills (NCS), frequently known as soft skills, are a key factor. Previous efforts to determine the recommended NCS for the ISD professional roles got limited results due to the use of manual procedures, absence of reference frameworks and small size of samples. This work presents a new approach which exploits existing reference models like ESCO, e-CF and NCSF (NCS Framework) as basis to analyse big datasets offered by tools like the ESCO database with thousands of relations between occupations and skills and the Ovate tool with millions of online vacancies. The combination of information from job market with the opinion of experts from ESCO provides a stronger basis for recommended NCS profiles for ISD professional roles, a promising option for an aspect not extensively explored.

Keywords: non-cognitive skills, soft skills, ESCO, e-CF, role profiles.

1. Introduction

The topic of soft skills and its relevance in information systems (IS) has attracted the attention of researchers and practitioners in recent years [16]. Current IS development (ISD) must respond to an environment where technologies are the drivers of digital transformation. This approach significantly influences business strategies, processes, products, and services by enabling new ways of working. This can only be achieved if IT professionals are equipped with technical skills but also with "non-cognitive skills" (NCS). The term NCS has been used by different authors [2], [11] to distinguish them from cognitive skills commonly measured by academic evaluation. However, other authors [6, 7], [14] prefer the term soft-skills over the one of NCS. We will use the term NCS in our research as it is the preferred term in documents of the European Commission [10].

As shown by multiple studies (e.g. [3], [13, 14]), the so-called non-cognitive skills are crucial in the success of educational performance, employability, income and professional development or career. There is a growing awareness that technical skills alone are insufficient for success in IT, particularly in today's dynamic, distributed and complex workplace [8]. At present time employers are less demanding technical skills to job candidates, because they are considering them trainable so they are prioritizing those who exhibit employability and NCS as positive attributes according to Newton et al. [17].

Albeit the evidence of the importance of NCS, researchers are finding several obstacles which hinder the progress in this field: absence of a common accepted terminology ranging from skills name to descriptions and lack of standardized or widely-accepted catalogue or model of skills. It is not strange to find many contributions both from literature (e.g. [14] compile up to 443 references) and from international reputed organisations (such as OECD, UNECO, WEF, P21, etc.) or from international projects given the relevance of the NCS. However, no solid common model has arisen from this big effort as, for example, there is very little consistency and uniformity in the existing models. Even more, many have not even explained in detail the rationale and methodology behind their development. Having a stable framework as reference and common language for NCS is a prerequisite to study the impact of these skills in professionals in IS development and operation. Such a framework would enable the analysis of information from different sources to determine the NCS profile demanded and recommended for IS profiles.

The everyday practice of ISD is very diverse as we can realise looking into ESCO, the European multilingual classification of Skills, Competences, Qualifications, and Occupations [9]. It describes, identifies, and classifies professional occupations (almost 3,000 for all sectors and around 100 for IT services). It also presents the recommended skills for them, selected from a global catalogue of more than 13,000. In the case of It profession, the European standard EN16234, known as e-CF [4], has also enabled the development of examples of role profiles for ICT professionals identifying area, competences and level of proficiency [5], a different point of view from ESCO. Other open data sources from job market (e.g. https://www.cedefop.europa.eu/en/data-visualisations/skills-online-vacancies) have shown a dynamic context characterized by fast evolution and big variety of professional profiles as well as they have confirmed the increasing demand of the IS related occupations. However, there is no clear results on the requested NCS in IS profiles neither strong agreement between the sources despite the existence of these references and the different studies mentioned in this research. Our contribution is aimed at providing a new approach to study the recommended NCS for IS profiles with stronger evidence than the previous contributions.

The structure of the paper starts with Section 2, which analyses existing contributions and summarises the findings and their limitations. Section 3 presents the different reference frameworks, which serve as basis for the analysis of NCS. Section 4 analyses a demand of NCS for IS profiles using open big data and data mining from different sources, discussing results and comparing them with conclusions from previous research works. The data for analysis is extracted from online vacancies (representing the present demand of NCS) as well as from models created by large groups of experts (representing trends and more prospective recommendations for profiles). The last Section summarizes conclusions and depicts further lines of work.

2. Traditional analysis of NCS in IS profiles

There has been little contribution from research on recommendations of NCS for ISD profiles in recent years. The existing studies clearly illustrate the strengths and weaknesses of the approaches adopted by the authors. The survey is a common method for collecting information in contributions like [12], [16] or [24]. The main problem of the survey method is the limitation in working with relevant samples and the possible inaccuracy of responses. In general, samples are limited in size and diversity of contexts: small samples such as 186 responses in [12] or 35 in [16], also combined with low responses rate such as 37.1% in [24] and data usually limited to one country. Moreover, authors of [12] and [16] do not refer to any existing framework to present the results as they create an ad-hoc classification. This is an additional weakness as there is not homogeneity in presentation of results, impeding the comparison with other studies or their use to populate a repository of comparable research results. However, one strong point is that personal experience of respondents can help to detect trends, but this is also possible by using big datasets.

Other contributions have followed another approach based on the analysis of the available from the job market. This is the case of the analysis in [20], [15] and [1]. All of them have manually analysed job advertisements to identify the most demanded NCS

related to ISD profiles with a personal classification based on authors' experience. The sample datasets are bigger than in surveys with more than 500 vacancies in each study. The results become more reliable and consistent, but the sample size remains insufficient to support relevant conclusions. Manual collection and analysis of job vacancies is time consuming and very limited to aspire to a proper coverage of reality. Moreover, the absence of a consistent framework for classification again obstructs the potential of aggregation of the different sources into a joint larger sample for a more solid analysis of data.

Our work is aimed at overcoming these limitations by adopting a new approach with two pillars: a reference framework for NCS based on accepted labour classifications like ESCO [9] and the exploitation of large datasets created by different initiatives of the European Union (EU). Our analysis will exploit the analysis of huge samples of job vacancies (millions of data, not a few hundreds) and the opinion of large numbers of experts devoting long dedication (hundreds during years).

3. Reference frameworks for analysis of NCS in IS development and operations

During the past decade EU has promoted several competence frameworks which could be applied to ICT profiles to enable a better coordination of the analysis of the job market. ESCO is a classification of occupations and their recommended profile of skills and knowledge. It uses a hierarchy of relationships between them as well as metadata and mappings to the International Standard Classification of Occupations (ISCO) to structure the occupations. Although they consider NCS in a way or another, they do not offer a specific framework structure by themselves. So, we will complement them with the NCS framework (NCSF) from the Skills Match project [19], which maps skills to the ESCO Classification and was developed after an exhaustive analysis of sources on NCS.

3.1. ESCO

ESCO is the European classification of skills and occupations. The aim of ESCO is to support job mobility across Europe and therefore a more integrated and efficient labour market, by offering a "common language" on occupations and skills that can be used by different stakeholders on employment and education and training topics. It provides descriptions for 2942 occupations and 13.485 skills linked to these occupations, translated into 27 languages. ESCO groups the occupations into ISCO-08 code groups with 4 digits classifying the different occupational profiles and showing the relationships between occupations, skills, competences. The ESCO classification was created in a process longer than 4 years by more than 200 experts from all the productive sectors, reviewed and refined by main stakeholders from education and training to labour market and authorities. According to implementation acts, the use of ESCO is compulsory for all Member States of EU from 2021 onwards.

3.2. Skills Match Framework

Skills Match was a project funded by European Commission (DG CONNECT) which developed and demonstrated a European-wide assessment and learning and guiding platform to help users to adapt their NCS to the demands of the labour market. The Skills Match project created a comprehensive and solid NCSF as the basis for its work [19]. The team analysed information from academic literature, reviewing 66 models and 403 publications with 2928 mentions to skills as well as 527 European Projects and other NCS existing frameworks and referenced models as OECD [18], P21 [21], UNESCO [22], WEF [23], among others. Skills Match uses ESCO as main reference and its 36 NCS of NCSF are mapped to ESCO skills generating 3138 connections at different levels. More than 700 buzzwords associated to each of the 36 NCS complement the description and help to identify mentions to NCS in other models and references. The visualization of the framework shows seven clusters for its 36 NCS, which group those ones most related among them (see Fig. 1).

Fig. 1. NCSF

4.   New approach for analysis of NCS in IS professional profiles

Our approach for the analysis of IS profiles covers two points of view: job market and opinion of experts. The market perspective extracts data from the analysis of job ads giving us a vision of recruiters who demand NCS to candidates. The expert perspective integrates their recommendation of NCS for each profile. Both visions can be combined to check differences and get stronger evidence of common elements. Although traditional analysis already used these two options, we have opted to dramatically increase the volume of information by exploiting open big data for the Skills-Ovate developed by CEDEFOP an agency of EU. Ovate extracts detailed information mapped to ESCO through natural language processing and machine learning on jobs and skills that employers demand in online job vacancies. This tool has collected data from July 2018 to September 2020 (and keep doing it) totalling more than 105 million of vacancies from thousands of sources (private job portals, public employment service portals, recruitment agencies, online newspapers, and employer websites) when writing this paper.
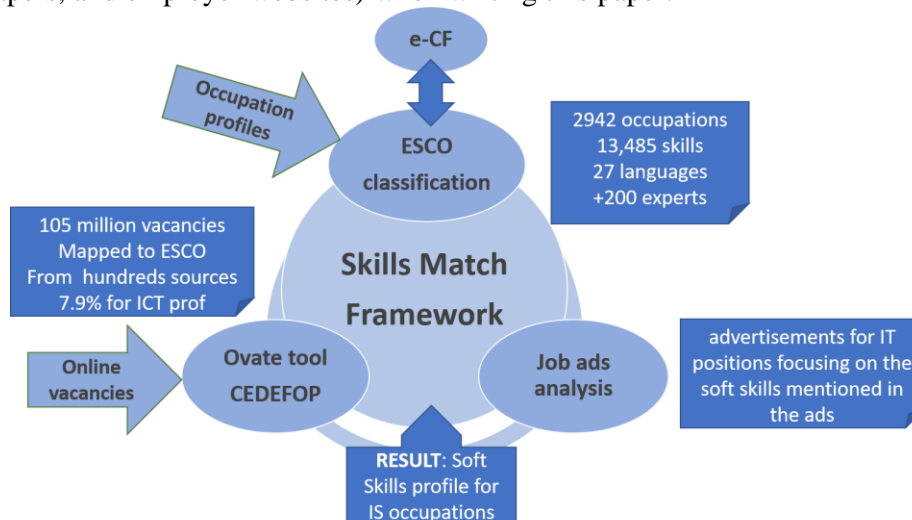

Fig. 2. Diagram for NCS demand analysis.

Fig. 2 shows our approach, which combines the sources of information with NCSF as central reference. The analysis in ESCO was possible because the authors created a local replica of ESCO database through the data downloading functions offered by the ESCO portal. It also shows a link to the standard EN16234-1 (e-CF) on IT professional e-

competences [4].

As an example of the potential of this approach, we have worked with the mentioned data sources for occupations of groups Software developers (subgroup 2512) and Systems analysts (subgroup 2511). These occupation roles were also manually analysed in [1] and [15]. Due to space reasons, we do not show results for other IS groups (e.g., 13 Technical managers and 35 ICT Technicians). In 2020, OVATE processed data for group of occupations 251 (Software and applications developers and analysts: 18 specialised occupations) included 2,804,416 OJA (Online Job Advertisements) in 27 European countries, 6.2% of OJA in all sectors. There is a clear difference with the 500 vacancies in [1] (which only detected 9 NCS), with 534 in [20] and with the 679 in [19] (which detected 17 NCS), all of them heterogeneous without reference to a NCS model. Thanks to the connection of ESCO to the NCSF, we can determine which the most mentioned NCS in OJA are along 2020, capturing big trends by recruiters and employers.

However, that one is not the only source of information with big amount of data. As commented, the ESCO database concentrates the work of hundreds of experts in linking almost 3,000 occupations to more than 13,000 skills and knowledge items with more than 114.000 relations. There are more than 100 occupations in the ICT services sector and, obviously, ESCO has recommended skills profiles for them. If we analyse the group 2511, there are 111 relations between NCS and the 18 occupations of the group while group 2512 has 18 relations for four occupations. We can compare this number of data with the 35 interviews of the study in [16], again with no reference to any NCS model. Table 1 shows the combined results of the NCS mentioned in job ads (as percentage of total ads where the NCS is mentioned) and recommended by ESCO experts for group 2511 and 2512 (percentage of occupations in the group which recommend the NCS).

Table 1. NCS recommended by ESCO experts and mentioned in OVATE's OJA for groups 2511 and 2512

	ESCO		OVATE			ESCO		OVATE	
NCS Name	**2511**	**2512**	**2511**	**2512**	**NCS Name**	**2511**	**2512**	**2511**	**2512**
accountability			4,0%	4,7%	motivate others	44,4%	50,0%	3,9%	0,9%
adaptability	16,7%	25,0%	3,7%	4,0%	networking	5,6%			
coaching	38,9%	50,0%	4,2%	8,0%	organisation			7,5%	8,5%
communication	72,2%	75,0%	19,8%	18,2%	personal development	27,8%	50,0%	1,6%	0,5%
conflict resolution	5,6%				problem-solving	50,0%	50,0%	8,9%	10,2%
critical thinking	22,2%		4,9%	1,0%	reliability	5,6%	25,0%	3,9%	3,8%
customer focus	55,6%	25,0%	1,6%		respect the environment	5,6%			
diligence			0,2%	4,5%	self-management	22,2%			
leadership	16,7%	25,0%	10,9%	16,3%	strategic thinking	5,6%		0,4%	
manage quality	44,4%		6,7%	1,4%	teamwork			14,0%	15,6%

5. Conclusions

Our approach shows how exploiting existing tools and sources with big volume of data (millions of vacancies in Ovate and thousands of relations between skills and occupations in ESCO) combining the expert view from ESCO and the market facts from OVATE provides solid results to the identification of recommended NCS for IS occupations. This approach is possible because there is a solid framework like NCSF linked to the ESCO classification. There is a big difference in using tens or hundreds of data in one single country in traditional manual studies to hundreds of thousands of records in Ovate or thousands of relations between skills and occupations in ESCO. We think that the exploitation of this amount of data with our approach provides more solid conclusions to the analysis of professional profiles in ISD. We have shown an example of the potential of analysis of NCS for specific IS profiles.

Adding the e-Competence Framework (e-CF) [4] with already 30 sample profiles based on e-CF competences [6] could strength our approach as it also represents a connection to

a standard adopted by big companies like Tata Steel, Airbus, Poste Italiane, etc. As ESCO and e-CF are also especially addressing technical skills and knowledge, it is also possible to work with them. However, we would need a previous non-trivial phase for correctly cataloguing specific skills within other more general concepts like competences in e-CF.

References

1. Ahmed, F., Capretz, L.F., Campbell, P.: Evaluating the demand for soft skills in software development. It Professional. 14 (1), 44–49 (2012)
2. Bowles, Samuel., Gintis, Herbert.: Schooling in capitalist America : educational reform and the contradictions of economic life / Samuel Bowles and Herbert Gintis. Basic Books New York (1976)
3. Brunello, G., Schlotter, M.: Non Cognitive Skills and Personality Traits: Labour Market Relevance and their Development in Education & Training Systems. (2011)
4. CEN: EN 16234-1:2019, e-Competence Framework (e-CF) - A common European Framework for ICT Professionals in all sectors - Part 1: Framework, (2019)
5. CEN Workshop on ICT Skills: CWA 16458-1:2018 European ICT Professional Role Profiles – Part 1: 30 ICT Profiles, (2018)
6. Clare Ignatowski: What Works in Soft Skills. Development for Youth Employment? A Donors' Perspective. Youth Employment Funders Group (2017)
7. Costin, G.P.: Legitimate Subjective Observation and the Evaluation of Soft Skills in the Workplace. Presented at the national training framework - training partnership and regional development, Albury, Australia (2002)
8. Damien Joseph, Soon Ang, Roger H. L. Chang, Sandra A. Slaughter: Practical intelligence in IT: Assessing soft skills of IT professionals. Communications of the ACM. 53 (2), 149–154 (2010)
9. ESCO Secretariat: ESCO Implementation Manual, (2019)
10. European Commission: Preparatory Action on Open Knowledge Technologies: Mapping and Validating Knowledge Call for proposals document. EUROPEAN COMMISSION, Brussels (2017)
11. Kautz, T., Heckman, J.J., Diris, R., ter Weel, B., Borghans, L.: Fostering and Measuring Skills: Improving Cognitive and Non-Cognitive Skills to Promote Lifetime Success. , Bonn (2014)
12. Lethbridge, T.C.: What Knowledge Is Important to a Software Professional? IEEE Computer. (2000)
13. Lindqvist, Erik, Westman, Roine: The Labor Market Returns to Cognitive and Noncognitive Ability: Evidence from the Swedish Enlistment. Research Institute of Industrial Economics (IFN), Stockholm (2009)
14. Lippman, L.H., Ryberg, R., Carney, R., Moore, K.A.: Key "Soft Skills" that Foster Youth Workforce Success: Toward a Consensus across Fields. Child Trends (2015)
15. Matturro, G.: Soft Skills in Software Engineering. CHASE. (2013)
16. Matturro, G., Raschetti, F., Fontán, C.: Soft Skills in Software Development Teams. IEEE/ACM 8th International Workshop on Cooperative and Human Aspects of Software Engineering. (2015)
17. Newton, B., Hurstfield, J., Miller, L., Page, R., Akroyd, K.: What employers look for when recruiting the unemployed and inactive: skills, characteristics and qualifications. Institute for Employment Studies on behalf of the Department for Work and Pensions (2005)
18. OECD: Definition and Selection of Competencies (DeSeCo). (2005)
19. Skills match consortium: Deliverable 2.1. Skills match consortium (2019)
20. Stevens, M., Norman, R.: Industry expectations of soft skills in IT graduates: A regional survey. The Australasian Computer Science Week Multiconference. (2016)
21. The Partnership for 21st Century Learning: P21 Framework Definitions. (2015)
22. UNESCO: UNESCO Competency framework. (2015)
23. World Economic Forum: The future of jobs, Employment, Skills and Workforce Strategy for the Fourth Industrial Revolution. (2016)
24. Wu, J.-H., Chen, Y.-C., Chang, J.: Critical IS professional activities andskills/knowledge: A perspective of IS managers. Elsevier. (Computers in Human Behavior), 2945–2965 (2006)