Intelligent analysis of digital economy competencies in agriculture labor

A A Aletdinova, M A Bakaev and V A Astapchuk
Automated Control Systems department, Novosibirsk State Technical University, Novosibirsk, Russia
E-mail: aletdinova@corp.nstu.ru

Abstract. As information becomes the main resource and production factor, there is widely declared need for digital competencies and skills in employees. This demand is uneven between industries and regions and is often biased due to companies’ and organizations’ desire to “keep up with the times”. In our paper we focus on employees’ digital economy competencies in agriculture, which is traditionally considered a rather low-tech sector, and compare them to several other industries with varying degrees of informatization. To negate the possible biases, in our analysis we rely on online data – vacancy ads published by companies at job-related portals – instead of traditional statistics. The online data was collected by our dedicated labor market monitoring intelligent system that has been in operation since 2011. The digital economy competencies were divided into 6 types: 1) intellectual and educational, 2) communicational, 3) motivational, 4) managerial, 5) ICT, and 6) ecological. We further use the groups and the wordings of the competences in online job-related ads to fill-in the “competencies specifications” thesaurus. Finally, we calculate frequencies for the digital economy competencies in vacancies for the different industries. We found that in agriculture the most demanded were the intellectual and educational, managerial and ICT-related ones.

1. Introduction
As the developed countries undergo transition towards digital economy, the demand for the related competences in workers is increasing in the labor market. There is now a term “critical competences” in use in the USA, similarly to the “critical materials” and “critical technologies” [1-3]. In our work, however, we employ a wider concept of digital economy-related competences (DECs), which correspond to the set of knowledge and skills most needed in the coming shift of the technological paradigm. Indeed, many researchers believe that “digital literacy” is the most essential competence in the XXI century [4, 5], while others register the corresponding synthesis of science, arts, engineering and design [6].

The American Occupational Therapy Association (AOTA) has developed standards for the “continuing competence”, which include such groups as knowledge, performance skills, interpersonal abilities, critical reasoning and ethical reasoning skills [7]. In [8], the authors identify computer literacy, media literacy, and media education as the key competences for the digital economy – and indeed, most research works related to DECs focus on ICT. However, we believe (somehow similarly e.g. to [4] who adds critical thinking to the set) that the set of skills essential for the new economy needs to be extended, based on the actual requirements by employers. Indeed, the World Economic Forum stated that by 2020 such competences should include complex problem-solving, critical thinking, creativity, managing and coordinating people, service orientation, evaluation and decision-making.
Higher education institutions need to make, negotiations, and cognitive flexibility. The Boston Consulting Group, based on their survey, identified such groups of DECs as cognitive, social-behavioural, and digital skills [9].

Understanding which competences will become truly critical should be based on expert evaluations, surveys, and monitoring and analysis of the actual labor market data. The latter approach has advantages in its natural capability to both register demand for DEC at a particular moment of time, and to study its changes with relation to the economic situation and the market conditions, even in retrospect. Since 2011, we have been collecting online data on job offers and ads posted by job-seekers (who mostly put online their resumes) for several regions of Russia. For this end, we employed the dedicated intelligent labor market analysis system that we developed and maintained [10]. The online data sources are popular Russian job-related websites, including the renowned HeadHunter.ru (http://headhunter.ru), which nowadays offers API for interested external parties and robust capabilities for analysis. The software’s database currently has more than 10 million unique records, structured per time periods, regions, industries and professions, etc. Previously, our analysis of the collected data was mostly focused on proposed and required wages and their relations. However, the data on job offers that we collect in the intelligent system also includes descriptions of job functions and skills, which call for semantic analysis with respect to the requirements towards DECs.

So, in our current work, we study the online data that is being collected by our software with respect to employers’ demand for DECs in such digitally “conservative” industry as Agriculture. In Section 2, we outline some particulars of online data mining, based on the experience from our labor market intelligent system project. We also list the DECs that were refined for the thesaurus subsequently used for the analysis of larger sample of job offers, collected by our software system. In Section 3, we present the results of the analysis of demand for DECs per different industries, with particular focus on Agriculture. In Conclusions we summarize our findings and outline perspectives for further research.

2. Methods and tools

2.1. Particulars of labor market online data mining.
The collection of online data can be performed through two principal approaches: either via API provided by the data source website (if available), or through web scraping of unstructured data from web pages. In the latter case, generally a virtual browser sends request to the website and returns the web page HTML code for processing.

For the web scraping-based collection, the challenges are changes frequently introduced in the source website’s web pages structure, and the web technologies that complicate the pure HTML-based data (such as AJAX, Web Components, etc.).

The necessity to monitor the rapidly changing labor market, with relatively short live span of an online job offer, calls for at least daily data collection. However, there’s the need to deal with repeated posting of the ads, which can be slightly modified by the owners.

The data gathered from online sources is often incomplete, both due to problems in its collection and to fundamental omissions. E.g. wages in job offers are not always stated, and if they are, only one border of the range is often supplied.

The target content, job-related ads and resumes, has weak formalism and possible spelling mistakes, which further complicates the data processing. Special methods of syntax and semantic analysis of the textual data need to be employed, often with increased computational costs.

This consideration together with generally large volumes of the collected data imposes requirements upon the storage system, involving not just relational databases, but also the modern NoSQL approaches.

Thesaurus is essential for structuring of collected data, so its population is the pre-requisite for any semantic data mining. With respect to our goals, it must include the concepts related to DECs, de-facto representing the conceptual model of the field involving job functions and requirements. The initial
population of the thesaurus necessitates both study of literature and the state-of-art data sources, and we describe the results of the respective analysis in the next sub-chapter.

2.2. Groups of digital economy competences.

From the analysis of existing sources [11–13] and based on the actual job offers published at HeadHunter.ru, we identified several major groups of competences related to digital economy:

1. **Intellectual and education-related DECs**, such as: Ability and readiness for life-long learning, Disposition towards improvement, Willingness to retrain, Professional mobility, Striving for new things, Critical thinking ability, Creativity, Self-development, self-improvement, Tendency towards non-standard solutions, Savviness, Resourcefulness, Systems thinking, Artistic creativity skills.

2. **Communicational DECs**, such as: Multi-language and multi-cultural understanding, Interdisciplinary communication, Customer focus.

3. **Motivational and will-related DECs**, such as: Disposition towards reasonable risk, Enterprise, Self-dependence, Self-realization, Striving for success, Self-respect.

4. **Managerial DECs**, such as: Project Management, Working with people, Willingness to work in a team and in a highly competitive environment, Working in conditions of uncertainty, Willingness to work both independently and in a group, Organized, Emotional intelligence.

5. **ICT-related DECs**, such as: Mastery of office programs, Data warehouses, Cybersecurity, Global data transmission networks and communication channels, Technical writing, Web development, Programming fundamentals, System administration and network configuration.

6. **Ecology-related DECs**, such as: Knowledge of requirements for waste disposal, Pollution control, Compliance with industrial hygiene, Safety and labor protection, Participation in resource saving.

7. Correspondingly, our thesaurus was populated with the above DECs and was used in the analysis of their frequencies in the published job offers.

3. Results

3.1. The establishing analysis (2018).

The employers’ demand for the DEC competencies in employees is specific per industries, so in our analysis sample we chose 5 diverse ones: 1) Agriculture, 2) ICT, 3) Health care and pharmaceuticals, 4) Mining, and 5) Financial sector. To characterize the demand, we measured “frequencies” – the share of job offers that mentioned at least one DEC from our thesaurus. The resulting frequencies extracted from the job offers we selected from HeadHunter.ru for December 2018 (106342 in total), are presented in table 1.

| Industry                     | Total job offers in the sample | Share of job offers with DECs, % |
|------------------------------|--------------------------------|----------------------------------|
| Agriculture                  | 4333                           | 68.2                             |
| ICT                          | 44358                          | 93.6                             |
| Health care and pharmaceuticals | 19682                          | 85.7                             |
| Mining                       | 4321                           | 51.6                             |
| Financial sector             | 33648                          | 88.9                             |
| Total/avg.                   | 106342                         | 77.6                             |

Since all the frequencies are above the 50%, it suggests that employers seek for employees with not just education and experience, as it was the case in “traditional” economy, but with DECs. The demand varies per industries, and our more detailed analysis discovered the following (Figure 1):

1. The intellectual and education-related DECs are in most demand in Agriculture and Health care and pharmaceuticals.
2. The communicational DECs are most asked for in Health care and pharmaceuticals.
3. The motivational DECs are most needed in Mining industry.
4. The managerial DECs are in most demand in Agriculture.
5. The ICT DECs are ubiquitous, and they are in nearly equally high demand in Financial sector, Mining, ICT, and Agriculture.
6. The demand for the ecology-related DECs are relatively low and was only found in Agriculture and Health care and pharmaceuticals, which suggests that this group is underestimated.

![Figure 1](image)

**Figure 1.** Comparative analysis of the demand for DECs per industries and the groups (1–6).

Among the considered industries, Agriculture is second to Mining in lowest demand for DECs (Mining actually had the least detailed job offers, 1/3 of them didn’t even have any requirements towards education or work experience). However, at the same time more than 2/3 of job offers in Agriculture require at least one DEC, and the proposed wages are not particularly low: e.g. 15000–140000 Rub for tractor drivers, 17000–45000 Rub for milking machine operators, 13000–250000 Rub for zoo technician, 12000–250000 Rub for agronomists (1 USD equaling approximately 65 Rub). Another particular feature is relatively fair demand for ecology-related DECs (in 10.02 % of the considered job offers), while for ICT, Mining or Financial sector this group is absent whatsoever.

### 3.2. The Additional Analysis (2019)

The vacancies data was repeatedly analyzed in May 2019, i.e. nearly 6 months after the first analysis. For each of the 5 selected industries, 100 job offers were randomly selected from HeadHunter.ru, with the scope set as all Russia (without specification of any particular regions). The work was performed by student workers who were given instructions replicating the establishing analysis. The results of the sampling analysis and the comparison with the 2018 are presented in table 2.

| Industry                        | Total job offers in search results | Share of job offers with DECs, % | Change vs. the 2018 data, % |
|---------------------------------|------------------------------------|----------------------------------|----------------------------|
| Agriculture                     | 5712                               | 81                               | +18.8                      |
| ICT                             | 60298                              | 84                               | -10.3                      |
| Health care and pharmaceuticals| 22083                              | 74                               | -13.7                      |
| Mining                          | 5384                               | 55                               | +6.6                       |
| Financial sector                | 33494                              | 85                               | -4.4                       |
| **Total/avg.**                  | **126971**                         | **75.8**                         | **-2.3**                   |
As the processed job offers were processed, the encountered DECs, as worded by the job posters (companies’ representatives) were included in the thesaurus that we are developing for further automation of the analysis. Each of the encountered \textit{DEC formulations} was linked to a Core DEC belonging to ones listed in Section 2 (see \textit{B. Groups of Digital Economy Competences}). In total, we selected 654 formulations – there was virtually no aggregation, since the goal was to have as many different workings as possible, for the sake of completeness of the planned automated analysis.

The thesaurus was implemented using CMS Drupal’s Content Construction Kit functionality. Examples of the Core DECs and of DEC formulations (for the popular Web Development competence) are presented in Figures 2 and 3 respectively (in Russian).

\begin{figure}[h]
\centering
\includegraphics[width=0.4\textwidth]{fig2.png}
\caption{Core DECs composed according to the identified groups.}
\end{figure}

The current version of the thesaurus is available at http://nstu.khvostov.ru/testcorolla/.

\begin{figure}[h]
\centering
\includegraphics[width=0.4\textwidth]{fig3.png}
\caption{DEC formulations extracted from job offers (for the Web Development competence).}
\end{figure}
4. Conclusions
The job-related websites providing online data can be used as indexes of labor market condition and the economic situation in general [11]. Particularly, in our work we used data collected from the popular HeadHunter.ru website to study the demand for DECs in several different industries, with particular focus on Agriculture. We devised the groups for the DECs, with the corresponding wordings of particular competences as specified in research works and found in the actual job-related postings. The thesaurus was used for semantic analysis of the large sample of data – more than 100 thousands of records for December 2018. After about 6 months we performed additional manual analysis of the data with respect to the DECs, using the same algorithm.

We found generally high demand for DECs (measured as share of job offers mentioning at least one DEC), as they were required in more than 50% of employers’ ads for all the considered industries. As one would expect, ICT and Financial sectors were the leaders, although in 2019 the former somehow lost the lead. The lowest demand was discovered for the Mining industry (51.61% in 2018 and 55% in 2019), while Agriculture considerably improved (from 68.18% in 2018 to already 81% in 2019). At the same time we should note that the two trailing industries gained the most during the 6-months period, while all the other industries lost in DEC share. It may suggest that there is the process of “regression to the mean”: after a while the degree of demand for DEC may become independent of a particular industry. It is likely that each employee would be expected to be in possession of up-to-date digital skills, just like in industrial society it has become customary to be able to read and write.

Further, we found that in Agriculture industry the employers did ask for DECs from the Intellectual, Managerial, ICT, and even the underestimated Ecological groups. Thus, even such conservative industry as Agriculture nowadays pushes employees towards self-development and self-learning, mastering systematic thinking, communication skills, project management, etc. Particularly impressive are the requirements towards computer literacy and computer professional skills – from competence in office software to network administration, programming, information systems development, and so on.

Our results suggest that the Agricultural industry is not left outside of the digital transformation and that the demand for DEC-skilled workers is becoming critical. Clearly, hiring of such employees can be possible only if the social infrastructure and the working conditions in rural areas improve considerably. We also expect that the currently underrated Ecology-related DECs may become crucial competences for the Agriculture in the near future, as the need for intensive and sustainable development of farming and animal husbandry becomes more prominent.

Our further research plans involve quantitative analysis of effects of particular DECs on wages in different industries and regions – e.g. how much more or possibly less money an employee possessing Technical writing competence would make in Agriculture. The millions of data records that have been accumulated in our labor market monitoring and analysis software system over the years should allow rather reliable estimations. They could be subsequently used by specialists in their career planning, as well as by HR specialists in planning job offers and seeking personnel with competencies that are critical for their companies’ businesses. Finally, educational institutions would be able to devise their study plans according to the current and prospective needs of the modernizing digital economy.

Acknowledgment
The reported study was funded by RFBR according to the research project No. 17-32-01087 a2 OGN. We would also like to thank the Bachelor and Master students of NSTU (Applied Informatics major) who contributed to the data collection: Daria Savchuk, Kirill Pyzhlakov, and Olga Tarsukova.

References
[1] Sousa M J 2014 International Journal of Economics, Commerce and Management II(9) 1–10
[2] Ala-Mutka K, Punie Y and Redecker C 2008 Digital competence for lifelong learning (Luxemburg: Office for Official Publications of the European Communities) pp 271–282
[3] Gallardo-Echenique E E, de Oliveira J M, Marquês-Molias L 2015 MERLOT Journal of Online...
Learning and Teaching 11 (1) 1–16

[4] Passey D, Shonfeld M, Appleby L 2018 Technology, Knowledge and Learning 23 (3) 425–439
[5] Hatlevik O E, Guðmundsdóttir G B and Loi M 2015 Computers & Education 81 345–353
[6] Walker I, Chan D K G, Nagami M and Bourguignon C (Eds) 2018 New Perspectives on the Development of Communicative and Related Competence in Foreign Language Education (Berlin: Walter de Gruyter GmbH & Co KG)
[7] Myers C T 2019 American Journal of Occupational Therapy 73 (2) 8
[8] Normore A T, Brooks J S and Silva SA 2016 Global Core Leadership Competencies: A Response to “Institutional Culture” and (In)-Competence in Higher Education The dark side of leadership: Identifying and overcoming unethical practice in organizations (UK: Emerald Group Publishing Limited) pp 155–176
[9] Russia 2025: from human resources to talents: the Boston Consulting Group 2017 report. Available at: http://image-src.bcg.com/Images/Skills_Outline_web_tcm27-175469.pdf
[10] Bakaev M, Avdeenko T 2016 Lecture Notes in Computer Science 9714 15–23
[11] Louvet J P 2000 Les principaux résultats de l'étude «Technologies clés 2005 (Ministère de l'Économie, des Finances et de l'Industrie)
[12] Alam K, Erdiaw-Kwasie M O, Shahiduzzaman M 2018 Journal of rural studies 60 60–69
[13] Voogt J, Erstad O, Dede C and Mishra P 2013 Journal of computer assisted learning 29 403–413