Using AI to Decrease Demand and Supply Mismatch in ITC Labour Market

Jussi Okkonen[1,*], Harri Ketamo[2], Hanna Lindsten[1], Teemu Rauhala[1],
Jarmo Viteli[1]

[1] Tampere University, Faculty of Information technologies and Communications Sciences,
FIN-33014 Tampere University, Finland
{jussi.okkonen, hanna.lindsten, teemu.rauhala, jarmo.viteli }@tuni.fi
[2] HeadAI, Rautatienpuistokatu 7, FI-28130 PORI, Finland
{harri.ketamo }@headai.com

Abstract. In information technology and communication (ITC) industry the technology advances are unexpected and moving mysterious ways causing significant mismatch between demand and supply in labour market. To some extent the mismatch is due to emergence of new technologies replacing the old ones. On the other hand, it is also due to lack of capability and capacity of educational system to provide up to date and spot on graduates. There are several attempts to bridge the gap between supply and demand, yet both public and private sector had at least partially failed in the task.

This paper presents a novel AI driven modus operandi that provides students or persons already in labour market a way to match one's competencies against existing or future competency requirements, and thus being valid employee or applicant. On the other hand, the service discussed in the paper is also tool for giving input for forecasting future labour needs. And the third, the service also serves as mid- and long-range planning apparatus for education provider when making decisions on what academic modules best serve the needs of society and individuals. The paper presents the concept of technology and functionalities as well as roadmap for developing and implementing the service. The presented concept is heuristically proofed in expert evaluation. Moreover, the heart of the service has been technically and functionally tested in several cases. At the conclusion the paper presents model for matching system for labour competency development. The issues related to implementation and the next steps are also discussed.

Keywords:: Labour market mismatch · Natural language· Digital twins · Competency development

1 Motivation for developing and researching the service

Extensive use of educational technology, especially digital learning environments, digital curricula, and digital managerial systems have brought about need for analytics

to monitor the use of the learning environments, the learning itself, and most important the individual competency development. Technology that is more sophisticated is provided by education technology industry to better serve teachers, individual, end users, as industries and primary customer. There are four main trends in enhancing teaching and learning practices. On micro level (learning event), analytics is for assessment of achieving certain goals. On meso level (subject level), i.e. implementing the curricula analytics are for achievements, adaptivity and general assessment. Macro level analytics is for promoting management by knowledge, risk assessment and measuring KPI's on different levels. The fourth trend is compliance, privacy, and security issues on the level mentioned and can be considered as the primary prerequisite for utilising learning analytics.[1]

In information technology and communication (ITC) industry the technology advances are unexpected and moving mysterious ways causing significant mismatch between demand and supply in labour market. To some extent the mismatch is due to emergence of new technologies replacing the old ones. On the other hand, it is also due to lack of capability and capacity of educational system to provide up to date and spot on graduates. There are several attempts to bridge the gap between supply and demand, yet both public and private sector had at least partially failed in the task. Universities are on the edge of education paradigm shift as labour market is not interested graduates with 3-year bachelor or 5-year master degrees, but they are actively seeking competencies to meet changing needs. Moreover, the shift from studies and labour is made more flexible and therefore also employees are actively seeking updates for their skillsets.

This paper presents a novel AI driven modus operandi that provides students or persons already in labour market a way to match one's competencies against existing or future competency requirements, and thus being valid employee or applicant. On the other hand, the service discussed in the paper is also tool for giving input for forecasting future labour needs. And the third, the service also serves as mid- and long-range planning apparatus for education provider when making decisions on what academic modules best serve the needs of society and individuals. The paper presents the concept of technology and functionalities as well as roadmap for developing and implementing the service. The presented concept is heuristically proofed in expert evaluation. Moreover, the heart of the service has been technically and functionally tested in several cases. At the conclusion the paper presents model for matching system for labour competency development. The service contributes practical work in education and competency development micro-level and macro-level.

2 Key functionalities and technology

The Service consist of three layers. The first user interface layer is for user management and access. The second layer is actual user interface for utilising service's analytical capacity, seeking for personal development scheme and putting input for analysis. The third layer is the data processing layer and the heart of the service. Figure 1 summarises the key functionalities and relations.
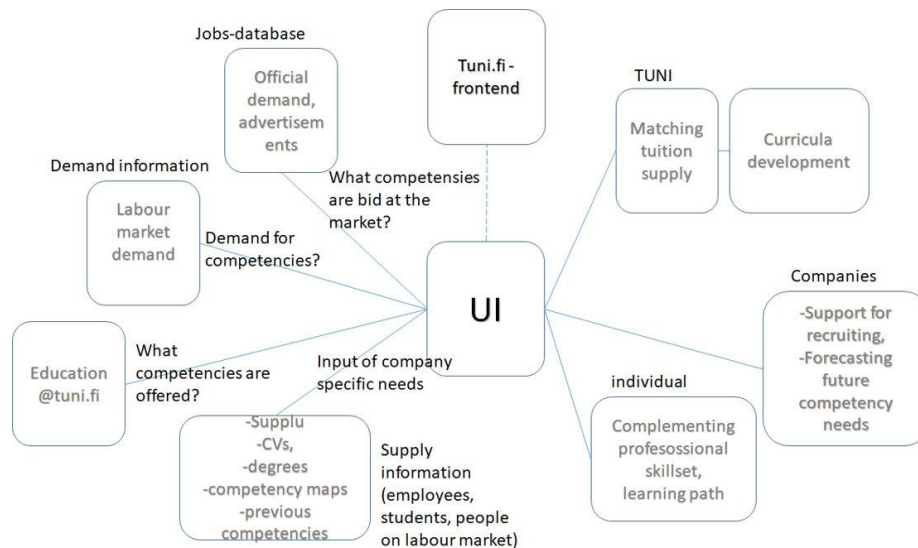
Figure 1: Outline of the service

Most of the natural language processing is currently done with machine learning algorithms focusing on multidimensional classification and/or grouping. That makes such applications narrow, dependent on training sets, and so makes cognitive reasoning, predictive analytics and explainable decision making very challenging. Furthermore, the privacy/security concerns related to black boxes are enormous.

Digital Twins are digital replicas on e.g. objects, entities or knowledge. Idea of Digital Twins raises from industry, but they have been applied in multiple contexts outside industry. Digital Twins serves as prepared and cleaned data layer that enables cognitive, predictive and explainable operations for next layer algorithms.

In this study, the strengths of unsupervised learning and reinforcement learning is brought together in order to build language based digital twins on knowledge domains. Digital Twins serves as prepared and cleaned data layer that enables cognitive, predictive and explainable operations for next layer algorithms.

Digital Twins on i) job demand, ii) person's current skills and iii) course offering are demonstrated as real world examples. Training data consist of +10000000 job openings, relevant news and curriculums/syllabuses. When also individual's competences are modelled as Digital Twin, algorithms can set up learning goals relevant for career. Because the course offering (online, onsite, P2P) is modelled AI and construct a learning pathway relevant for career. [2,3]

From company point of view Digital Twins can be used to strengthening both career planning and recruiting pipeline. In national level Digital Twins help industries to get optimal talent (minimum training time) and training companies / universities to build optimal curriculums (for companies and individuals). Figure 2 presents general outline of AI powered core technology.
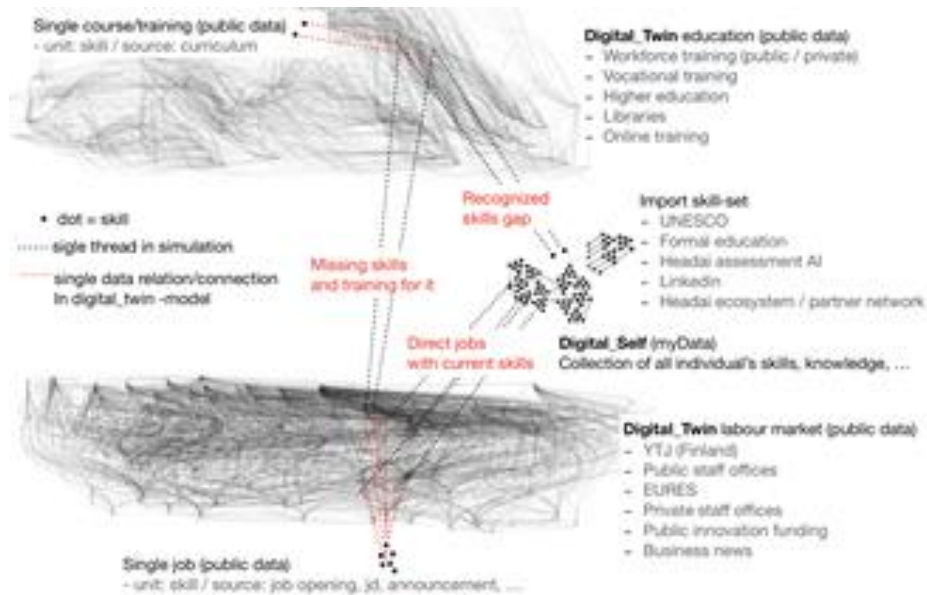
Figure 2: Outline of the core technology

In terms of anonymising data [4], digital twins can be used to ensure only meaningful words will be included into anonymised data. In other words, we construct Digital Twin on medical record by using e.g. general ontology and medical ontology in selected language. Such Digital Twin will be based on only meaningful set of verbs and substantives (including their derivatives and compound forms if meaningful) and so basically every sentence is 'subject-predicative-object' type of sentence. No identities will be passed to Digital Twin, because general ontology and medical ontology do not have general identities inside the model. However, some identities, like 'Alfred Nobel', are there, but such identities' meaning is based on compound form of the word and no 'Alfred' nor no 'Nobel' will be passed to model. In general, if identity is not in encyclopedia, it is safe.

There is naturally trade off. In this case trade off is the details. After Digital Twin on medical record is built, lot of details that raw information had, is cleaned out. Finally, if we need to make maximum safe anonymised data, at least two different methods should be used. By combining Headai Digital Twin -model and traditional supervised machine learning anonymization we can be sure we have more secure data that with only one method.

In general, even though in national and regional level we should consider skills as economical data, in individual level it is confidential information like health records. Not necessarily the skills person has, but what the person is missing compared to group averages. That makes anonymization crucial in this kind of approaches. To enable economical analysis without risking individuals' privacy.

As stated above, the service utilises novel way to approach competency development on individual and company level. At this point the feasibility of the technology is

already known and the actual implementation will start. The following section presents the roadmap for implementing the service.

3   Roadmap for building the service

The service is implemented to Tampere University web environment and the user interface follows the graphical appearance of any other service provided by the university. However, the actual user interface is adaptation of Microcompetencies by HeadAI and it is individual setup to university context. This step also consists of reading supply of academic modules from the university databases.

The second phase of the implementation is to build individual profiles for any user. This requires separate software component and it is for reading data of course completition from the university database as well as other relevant data on completed degrees or professional development courses. In addition to that the consent by individual for utilising data possessd by university is also needed at this point. This considers especially alumni, since there is already existing body of data based on their prior activities.

The third phase is the mapping, recognition and definition of skills needed by the companies. These are also translated to educational needs and operated into academic modules. It starts with mapping the demand for labour and acknowledging competency gaps. Demand information brings about the critical competencies and shed light on the rivalry on competent employees. On the other hand, there should be also mapping on the supply. Especially what is delivered by the educational system, i.e. what academic and professional development modules are available. The mismatch is visualised in the user interface for supporting decision making.

The fourth phase is to enable company specific competency models for creating development schemes. This is done by refining the data provided by the companies. Also, the stakeholder data is utilised in this phase. The refinition is based on text form data entered to system. This data will be also augmented by data gathers from national sources consisting of for example trend analyses, macro level news, and investment data. The augmentation enables forecasting development trends and foresee the change in needed academic modules.

The fifth phase is screening and picking academic modules to meet company requirements as well as individual needs. It depends on the user what modules, programs or professional development schemes are highlighted. Individuals wil get more precise recommendations to match their competency gaps, yet companies receive general recommendations according to their own development avenue. The competency development is based on the topics university already provides.

The sixth phase is composing development scheme for complementing academic modules. This phase is data driven and it is extension of the labour market demand analysis executed before. This requires also dialogue with existing supply of academic modules. This functionality is aimed for the developing education and can be considered as a positive externality of running the service.

The seventh phase in implementing the service is to provide individuals personalised analysis on their labour market fit. This applies to all individual using the service, despite if they are students, alumni, or external users. By the personalised analysis

one's labour market position is recognised and personal competency development scheme is outlined. By aggregating this data, a holistic view on labour supply and demand is made.

The service provides for individual, company, industry and university. The motivation for individual to participate and surrender personal data is based on accurate professional development or even early stage guidance to specialise in the studies. Companies benefit of complementing their company specific competency development by having reflected their position to general development of the industry as well as seeing the feasibility of new technologies labourwise. The industry benefits for bridging the competency mismatch gap and proving educational system input on future needs for academic modules. For university this provides critical market information and helps in development activities of education.

4   Discussion

The presented service is based on design and heuristic testing of concept. It will be implemented during spring 2019 and it is fully operational in spring 2020. Since the technology utilised in the implementation is existing and requires only little tailoring to meet the requirements the further work is about validation of the concept, operation model and arranging the setup for meeting external requirements. There are at least two major issue to be contemplated. The first issue is the ethical considerations about using the service, and the second one is the validation of the service.

The ethical considerations are issue on the perspective of individuals, and on the other hand on the perspective of the labour policy. The individual perspective is about privacy and security as well as maintaining control on individual. The privacy is quite easily tackled if there is responsible service provider and sufficient guidelines to use the services. The security is issue as it is with any data repository containing personal information. The control issue is the most complex one, since using the service individual surrenders control, at least to some extent. The control on self-regulation and decision power on own career is in flux. Moreover, the revealed competency gaps are efficiently recognized and communicated also to employer.

The labour policy perspective is not straight forward either. By using service to analyse the trends and setting development agendas it makes decision making easier, yet it might make long range planning more difficult as it might attract to offer academic modules that serve short-term needs. However, this is the issue of sound judgement and therefore beyond the scope of the development and research endeavour.

The validation of the service is issue of several open questions. The first one is that is a (simple) market test sufficient way to validate the concept? As brought about previously, the technical issues are solved and now the implementation is more about integration and setting the operational limits. This tends to suggest that concept is already validated by a market test and assessment should go further. The next step is user experience research in wild with various user groups. Most evident are companies, students, and administrative users. The second tier of users are the key stakeholders, yet their insights are needed too. What is needed along the user experience research is the proof for validity of actions, say assessment of the effect on labour market, effect on individual's labour market position, assesment of competency de-

velopment in industry, and the effect on development of industry locally. This is derived straight forward the presumed cause and effect justification, yet there are several sub-studies to be conducted. In near future the next steps will consist of implementing and testing the service as well as recruiting users.

References

[1] Okkonen, J., Helle, T., Lindsten, H. (2020) *Expectation differences between students and staff of using learning analytics in Finnish universities*. In Proceedings of ICITS2020

[2] Ketamo, H., Passi-Rauste, A., Vesterbacka, P. & Vahtivuori-Hänninen, S. (2018). *Accelerating the Nation: Applying AI to scout individual and organisational human capita*l. In proceedings of ICIE2018 International Conference on Innovation and Entrepreneurship, March 5th-6th 2018, Washington DC.

[3] Ketamo, H., Moisio, A., Passi-Rauste, A. & Alamäki, A. (2019). *Mapping the Future Curriculum: Adopting Artificial Intelligence and Analytics in Forecasting Competence Needs*. In Sargiacom, M. (Ed.) Proceedings of the 10th European Conference on Intangibles and Intellectual Capital ECIIC 2019, Italy, 23-24 May, pp. 144-153. ISBN: 978-1-912764-19-8 ISSN: 2049-0941

[4] Alamäki, A., Aunimo, L., Ketamo, H. & Parvinen, L. (2019). *Interactive Machine Learning: Managing Information Richness in Highly Anonymized Conversation Data. In L.M*. Camarinha-Matos, H. Afsarmanesh & D. Antonelli (Eds.), Collaborative Networks and Digital Transformation. The Proceeding of 20th IFIP WG 5.5 Working Conference on Virtual Enterprises, PRO-VE 2019, pp. 173-183