EUROPEAN
POLICYBRIEF

Decision support for policy makers: An intelligent system with coherent knowledge of diverse lifelong learning interventions in EU countries

Rong Qu

INTRODUCTION

Among lifelong learning researchers and practitioners, there is rich practical knowledge of what makes educational interventions for young people successful. This provides sound foundations for building an Intelligent Decision Support System (IDSS) to support informed decision making in policy making. There is, however, a lack of research spanning the two disciplines – Education and Computer Science – which underpin building an IDSS for this purpose. This policy brief reports outcomes from an interdisciplinary project sponsored by the European Commission's Horizon 2020 research programme.

Based on novel research conducted on knowledge representation and reasoning, an open interactive demonstration ENLIVEN IDSS (http://enliven.cs.nott.ac.uk) has been built to demonstrate the feasibility of IDSS to support policy making in education and training in Europe. With more than 200 interventions for young people Not in Education, Employment and Training (NEET) in its knowledge base, the ENLIVEN IDSS can retrieve interventions similar to queries made by users via

an interactive interface in different languages. Novel data mining techniques have been applied to explore machine learning in this new domain. Recommendations are made on addressing challenges and future research directions.

This work was carried out in Work Packages 8 and 9 of the ENLIVEN project, led by Computer Science, The University of Nottingham.

EVIDENCE AND ANALYSIS

The foci of this research are:

- First, a unified template of attributes (i.e. knowledge representation) was built to represent various interventions for young 'NEET' people across European countries as cases in a consistent form. A similarity measure model was built to assess how similar cases are based on key attributes.
- Second, a basic IDSS was built with the collected cases stored in a consistent template (see Figure 1). When a stakeholder or policy maker raises a query, the IDSS retrieves stored previous cases, ordering them by their similarity to the query to support more informed decision making.
- Third, an online interactive ENLIVEN IDSS (http://enliven.cs.nott.ac.uk, see Figure 2) has been built to retrieve similar cases (see Figure 3). It also supports advanced analysis using statistical tools and information theory. The interactive ENLIVEN IDSS has been recursively enhanced based on feedback collected from different stakeholders and potential users.
- Finally, advanced machine learning has been applied using several data mining and text mining techniques to extract interesting knowledge from the ENLIVEN IDSS across all the cases collected.
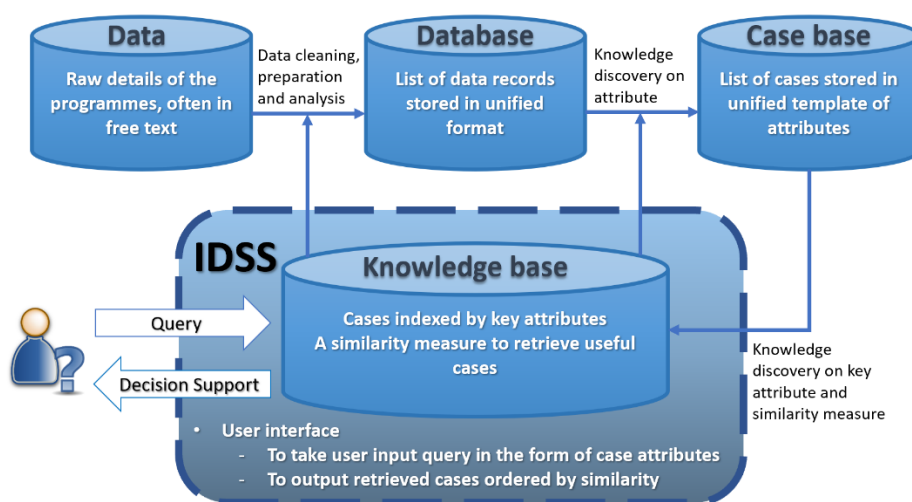

Figure 1. The development of ENLIVEN IDSS

Figure 2. The interactive user interface for ENLIVEN IDSS: query page

Figure 3. The interactive user interface for ENLIVEN IDSS: retrieved cases

Figure 4. The interactive user interface for ENLIVEN IDSS: adding new cases
(Note: only the system administrator can add new cases to the case base)

The interactive demonstration IDSS has been built in English, French, and Chinese versions, with the representative NEET interventions as an area of policy focus. It can be further extended by an administrator with new cases for NEETs (see Figure 4). These can be collected over time from the literature, from policy makers or stakeholders, and stored in the same unified template. It can also be extended with cases and interactive user interfaces in more languages. With the same methodologies, other IDSSs can be built with interventions for other target groups in lifelong learning.

Details of the first and second research outcomes mentioned above are reported in ENLIVEN Policy Brief No. 3. The third and fourth research outcomes are summarised here. Recommendations to address challenges and future research directions identified are presented in the final section.

Research finding A. With the established unified case template, further data analysis on the 222 cases collected revealed that the case attributes are not well distributed.

To facilitate effective documentation of rich knowledge in policy making practice, a unified template with 78 structured attributes has been built, storing 222 NEET interventions as cases in a basic IDSS. This allows the IDSS to distinguish cases from one another, and to query them in an effective manner. As detailed in ENLIVEN Policy Brief No. 3, "at present documents take diverse forms and often lack sufficient detail" to support cross-referencing, comparisons, and thus informed decision-making. The policy documents collected are of variable type (pdf, Word, web sites, databases), length (from two-page posters to hundreds of pages of analysis) and level of expertise (depending on the author).

Four key attributes (Locations, Target groups, Aims and Activities) in the case template have been selected to calculate the similarity between cases. For each of these key attributes, categories were extracted as their possible values in different cases. Statistical tools and information theory were used to conduct more analysis on the key attributes.

Using statistical analysis, we found that the categories for the key attribute "Target Groups" have a relatively balanced distribution across all 222 collected cases; however, the other three key attributes "Activity", "Aims" and "Location" are less balanced. This presents a potential issue for effective comparisons between cases, and thus for retrieval of useful cases from the IDSS. Further statistical analysis on grouped categories of the key attributes indicated that an improved balance of the categories as the values for the key attributes could be achieved, especially for "Aims" and "Activities". This justifies and supports the grouped categories selected by human experts to calculate the similarity between cases in the ENLIVEN IDSS.

Using information theory, the four key attributes were further analysed using information entropy, which measures the quantity of surprises (i.e. uncertainties) generated from a collection of possible outcomes. The four key attributes all have a reasonable amount of entropy, which again justifies their selection for calculating similarity between cases.

Research finding B. Machine learning can extract and reveal interesting knowledge from cases collected, and process it using different text mining techniques.

The ENLIVEN IDSS has been built with an interactive user interface and enhanced based on feedback collected from different potential users. In addition to English, the user interface could be switched easily to two other languages (French and Chinese – where the Chinese version uses a very different encoding for demonstration purposes).

After data preprocessing, which cleans and transfers the cases into a vector of words, several data mining techniques (concept extraction, concept categorisation, concept clustering and association rule extraction) were used to extract hidden knowledge from the 222 cases in the ENLIVEN IDSS. One example of automatically extracted knowledge using the associate rule technique is as follows:

> The results imply that UK interventions are aimed at people with disability and health problems whereas those in Australia focus on offering basic skills in literacy and numeracy. A possible explanation for this difference is that in the UK the proportion of 16-24 year olds who were NEET has been found to be higher for those with disabilities (29%) than those without (9%) (Powell, 2018). In Australia, more than 30% of young people who left school at Year 9 or below, and thus lack basic education, were found to be NEET in 2012 (Wong, 2016). However, a significant number of NEETs in the UK have also been found to be lacking in basic education. In the UK 23% of 16-24 year olds without any qualifications were found to be NEET compared to 9% amongst those with a qualification of level 2 and above (Powell, 2018).

The ENLIVEN IDSS we have developed, with its unified representation of interventions, presents an excellent platform for future advanced research and for research communities to extract interesting hidden knowledge using other artificial intelligence and machine learning techniques.

POLICY IMPLICATIONS AND RECOMMENDATIONS

This ENLIVEN research generated interesting findings, leading to the following recommendations and suggested future research directions on developing computer-aided decision support systems:

Recommendation A. A consistent representation or a domain taxonomy is highly desirable in building future intelligent systems in policy making in both research and practice.

Along with the rapid development of intelligent systems across many applications comes the pressing demand to collect large amounts of data and represent domain knowledge in coherent and efficient ways for research and practice. The demonstration ENLIVEN IDSS has been built based on a unified case template which models various NEET interventions in policy making in lifelong learning; otherwise cross-referencing and comparisons between interventions to learn from lessons, reuse good practice and support informed decision making would not be possible. To build future intelligent systems in policy-making for other programmes with different target groups or aims across different countries, and thus facilitate effective and efficient decision-making, the same methodology could be adopted. This requires a standard for data recording at the EU level to encode and represent knowledge from practitioners and policy makers.

In information science, a domain taxonomy is a representation and definition of a set of categories, properties and relations between concepts and entities that are of importance in a particular domain. For example, the biomedical literature uses numerous taxonomies and ontologies to enable efficient automated processing of literature of various subdomains (Konopka, 2015). In order to improve the automated processing of the documentation of interventions, a domain taxonomy in lifelong learning would be of great use, providing a controlled vocabulary and allowing for a consistent description and modelling of interventions. This presents a challenge in both practice and research, across education and computer science. However, it would be highly valuable for future documentation in policy-making and the development of intelligent computerised systems.

Recommendation B. Consistent evaluation approaches are crucial to provide quantifiable metrics of success when comparing different interventions, and support building informed and automated systems.

Although the feedback collected from sampled users on the interactive ENLIVEN IDSS is in general positive, there is not enough information in the current literature to provide clear and commonly agreed formal measures of success. By quantifying previous success, the ENLIVEN IDSS can advise both on a specific set of previous interventions, and against approaches that it knows were not successful (or at least issue warnings that attention should be directed towards that intervention, in order not to repeat past mistakes).

For a complex domain such as policy-making in lifelong learning, with different aims for various target groups across different countries, there is unlikely to be a universally applicable evaluation approach. However, it would be valuable to explore the scope for such evaluation approaches across different types of interventions using similar metrics based on consistent vocabularies and taxonomies. This would allow automated systems to infer potential reasons for success or failure, and to improve over time, by helping human users of the system differentiate good practice from bad.

References

Konopka, B. M. (2015). Biomedical ontologies—A review. *Biocybernetics and Biomedical Engineering*, 35(2), 75-86.

Powell, A. (2018). NEET: Young People Not in Education, Employment or Training Briefing Paper Number SN 06705, 24 August 2018, House of Commons Library. https://researchbriefings.parliament.uk/ResearchBriefing/Summary/SN06705 (accessed 19 Feb 2020).

Wong, M. (2016). The characteristics of youth not in education, employment or training (NEET) in Australia and Japan, Japan Institute for Labour Policy and Training, Tokyo http://www.jil.go.jp/profile/documents/m.wong.pdf (accessed 19 Feb 2020).

RESEARCH PARAMETERS

The ENLIVEN research models how policy interventions in adult education can become more effective. Different work packages focus on the role of governance and policy, participation, workplace learning and adults' well-being. It implements and evaluates an innovative Intelligent Decision Support System and provides a new and more scientific underpinning for policy debate and decision-making on adult learning, especially for young adults. The project investigates these lifelong learning aspects through quantitative and qualitative analyses.

PROJECT NAME	Encouraging Lifelong Learning for an Inclusive & Vibrant Europe (ENLIVEN)
COORDINATOR	Professor John Holford University of Nottingham, Nottingham, England, United Kingdom john.holford@nottingham.ac.uk
CONSORTIUM	University of Nottingham – Nottingham, England, United Kingdom 3s Unternehmensberatung GmbH – Vienna, Austria Bulgarian Academy of Sciences, Institute for the Study of Societies and Knowledge – Sofia, Bulgaria KU Leuven/University of Leuven, Leuven – Belgium Slovak Academy of Sciences, Centre of Social and Psychological Sciences – Bratislava, Slovakia Tallinn University/Tallinna Ülikool –Tallinn, Estonia Universidad de Deusto – Bilbao, Spain University of Edinburgh – Edinburgh, Scotland, United Kingdom University of Melbourne – Melbourne, Australia University of Verona/Università degli Studi di Verona – Verona, Italy
FUNDING SCHEME	European Union Horizon 2020 Framework Programme for Research and Innovation (2014-2020) – Societal Challenge 6 – Europe in a changing world: inclusive, innovative and reflective societies", call YOUNG-3-2015, topic "Encouraging Lifelong Learning for an Inclusive and Vibrant Europe (ENLIVEN)" Grant Agreement No. 693989
DURATION	October 2016 – September 2019 (36 months).
BUDGET	EU contribution: €2 499 788.50
WEBSITE	https://h2020enliven.org/
FOR MORE INFORMATION	Contact: Professor John Holford john.holford@nottingham.ac.uk Contact: Ruth Elmer ruth.elmer@nottingham.ac.uk
FURTHER READING	*Related ENLIVEN publications:* Clos, J. & Palmer, C. (2020) *Data Mining in the ENLIVEN Project.* ENLIVEN Report D8.2. Clos, J. (2019) The Architecture of the ENLIVEN Intelligent Decision Support System. ENLIVEN Report D9.1. Palmer C. & Clancy S. (2019) "Computer-aided Decision Support to enable policy making in Lifelong Learning". Forthcoming in Holford, J., et al (eds), *Lifelong Learning, Young Adults & the Challenges of Disadvantage in Europe*, Basingstoke: Palgrave Macmillan.