

Deliverable Report

The logo for COMPLEAP, featuring the word "COMPLEAP" in bold, black, uppercase letters on a yellow, trapezoidal background.

Agreement number: SI2.488704 (ECOKT2016-1)

Project name: Learner-centered digital ecosystem of competence development (CompLeap)

Date: 31.5.2019

Work package: WP3

Title: D26 Three Prototypes

Author: EDUFI

Description of the Action:

EDUFI has from February to May 2019 produced three interoperable and interlinked modular prototypes to prove the practical functionality of the CompLeap framework. The modules consist of:

- Module 1: Local study record service integration (past)
- Module 2: Competence profile with current competencies (present)
- Module 3: Suggestions for educational opportunities (future)

Outcome of the Action:

As a result of the project, three interoperable prototype modules have been developed. The proof of concept prototypes can be accessed through this link:

<https://poc.compleap.testiopintopolku.fi/>

In the service the user can:

- see her previous education
- fill in her previous education from abroad
- tell her own future interests
- get recommendations of interesting study options based on her previous education and future interests
- get recommendations of interesting study options based on her future interests

These prototypes have been developed in the EDUFI test environment.

Finally, tests relating to creating analytics functionalities are described.

Table of Contents

The Three Prototypes in the ComLeap Project.....	1
Background.....	1
Aim of the Prototypes	1
Implementation of the ComLeap Prototypes	2
Implemented by the End of May 2019	2
Technical Implementation	3
Further Development.....	3
Data Sources and Restrictions of Beta-prototype	3
Prototype Content from a Learner's Perspective	4
Module 1: Prior education	4
Module 2: Current interests.....	5
Module 3: Learning Opportunity Recommendations (Analytics Prototype).....	6
Testing and Developing Additional Analytics Features.....	8
Background.....	8
Analytics test nr.1: Condensation of the ESCO Competencies to be Used as an Interest.....	9
Results	9
Action Taken.....	9
Suggestions.....	10
Analytics test nr.2: Matching eRequirements Degree Unit Descriptions to Two ESCO Pillars.....	10
Encountered problems.....	11
Approaches to solve these problems.....	11
Results	12
Summary of the results	12
References.....	12

The Three Prototypes in the CompLeap Project

Background

To obtain a full picture of the CompLeap prototypes, this following document is complemented by Deliverable 24 Open source code for all components and Deliverable 25 Technical Documentation, including glossary, model and architecture. A full description of the user flows within the prototypes can be found in Deliverable No 21: Feedback and specifications to user scenarios in WP2.

The principle objective of WP3 was **to develop three prototypes and to prove their practical use** under real-world conditions. This document describes the content and implementation of the CompLeap prototype application. A brief overview of the general concepts behind the prototype is offered first. Following that, the core application components are described in more detail.

The core WP3 Tasks included:

- T1.1: Setting up test environment including suitable test dataset to be used in WP 3 and WP 4
- T1.2: Development of prototypes taking into account of learner perspective
- T1.3: UI and Business logic to support the targeted learner-centred services (personalized learner path, continuous applying)

Within WP3, the lead of the work package, Finnish National Agency for Education (EDUFI), produced testable prototypes that consists of **an interface, content and associated services**. Additionally, the University of Oulu has been testing out various methods of integrating AI and analytics functions for the benefit of the learner within the existing prototypes.

Aim of the Prototypes

The joint prototype application demonstrates a learner profile that makes use of information about learner's skills, competences, and interests. With the aid of this information, learning opportunity recommendations are offered to help the learner realise new areas where the competences and interests could be combined.

On a logical level, the application consists of three interlinked modules: prior education and competences, current interests and aspirations, and recommendations for finding new ways to combine these. While each of the modules demonstrates a particular viewpoint to CompLeap, the modules also work in tandem to form a learner profile for the user.

The prototypes **give an understanding of the applicability of the final product, as devised and planned in WP2**. The developed prototypes have been developed with associate partners, i.e. education providers. During the onsite testing phase, the partners operate, evaluate and optimize the prototype. The development process has been undertaken in close cooperation with the

University of Oulu, and with DUO in the Netherlands. University of Oulu has been in charge of testing various Natural Language Processing solutions for the service, whereas DUO has been researching the possibilities of uptake in the Netherlands. The final product is aimed to be implemented at least in Finland and in the Netherlands.

As the proposal builds on a learner-centered, user-driven, development approach, the project has worked with user organisations for development, iterations and real-life testing of prototypes through its partners and reference groups. The results of these tests have been used to refine the prototype. The testing so far has led up to the actual pilot deployment in the WP4.

The final product will be a service that can be integrated into another service or service framework, and as such the aim has not been to create a separate CompLeap service to be used on its own. The aim here has been to build upon existing trusted services, create added value for service providers and for users, and gain possibilities for wider dissemination.

The testing will aim to involve different user groups from young adults to immigrants and refugees. In addition, number of interested reference group members have expressed their interest to participate in the development and testing.

Individual's past achievements such as certificates will be included electronically in machine-readable form to the self-assessment. This needs to be defined further. Perhaps through machine-readable Europass content.

Implementation of the CompLeap Prototypes

The CompLeap prototypes are being developed in two sessions. The primary phase was from February to May 2019. During this phase the test environment and test datasets have been created. The piloting has begun in May 2019, and the piloting process will feed into the development in the second phase until the end of August 2019.

Implemented by the End of May 2019

Competence Profile

- The learner's competencies are shown to the learner based on their local study record data and data from eRequirements (available in Finland)
- The learner can provide prior and future interests

Recommended educational offer

- based on interests and past studies

Technical Implementation

Technologies	JavaScript, React, R, AWS (infra)
Release notes	https://jira.oph.ware.fi/jira/projects/CL?selectedItem=com.atlassian.jira.jira-projects-plugin:release-page
Bugs	https://jira.oph.ware.fi/jira/secure/RapidBoard.jspa?rapidView=163
Link to service	https://poc.compleap.testiopintopolku.fi/
Licence	EUPL

Further Development

The user piloting of the test data and prototypes began in May 2019. Improvements to usability and tweaks based on feedback received in piloting are implemented by the 30th of August.

Data Sources and Restrictions of Beta-prototype

The first implementation case for the beta prototype is in the Finnish Context. Under the current Finnish legislation, the Beta-prototype cannot utilize real personal data, even though we as a National Agency for Education in Finland have access to data sources with reference to the KOSKI-service. For this reason, we employ pre-confined data for the profiles of the test users in the Beta-prototype.

In practice, this signifies that the test users cannot log in and authenticate to the prototype with their own personal data and thus visualise their past study records in the Competence profile. The data of the test users represent some of the most general user groups identified in the CompLeap user research. The most potential user groups and personas, and thus the most fruitful user profiles were identified as follows:

- 1) user with full vocational degree in environmental or arts domain
- 2) user with discontinued education in technical or health care domain
- 3) user with foreign education

The prototype will nonetheless utilise the vocational qualification data of the Finnish national electronic service for educational qualifications, eRequirements service. Data from vocational educational qualifications is visualised in the Competence profiles of the test users.

In the CompLeap beta-prototype, other levels of educations in relation to the user's previous study records have been framed out of the final Beta-prototype and its modules. The decision to primarily start off with the VET sector was done due to good availability of VET qualification requirements set at the national level of Finland. Thus the availability of suitable data sources was one of the key argument select users with vocational study background.

In addition to this, the prototype includes a direct interface to the educational offer. Educational recommendations utilise current vocational educational offer from the national studyinfo.fi service.

Prototype Content from a Learner's Perspective

The three prototypes planned in [Annex 1](#) are embedded in the Learner Plan prototype and its three modules:

1. Local study record service integration (past)
2. Competence profile with current competencies (present)
3. Suggestions for educational opportunities (future)

The key aim was to create a link between the competence profile (including interests and existing competencies) and the provision of educational opportunities. Compared to prior phases, this phase includes real competencies as well as proper educational offers. The utilization of national (educational) databases within the service will be tested and illustrated.

Module 1: Prior education

The prototype for prior education demonstrates two different scenarios: verified education and non-verified education. The distinction between the two is that in the verified education scenario, a central, authoritative repository of study records that contains information about learners' prior education is required, whereas in the non-verified scenario this requirement is removed.

In the prototype module 1, **KOSKI**, Finnish national register, is used as a demonstrative example of a provider of verified education information. KOSKI register contains study records for verified education done in Finland for multiple different levels of education. These records can be used together with textual descriptions of the studies (i.e. the eRequirements in the Finnish case example), which offers a way to examine learner's competences. In the prototype, the usage of the repository is demonstrated with a mock implementation and test learner profiles that correspond to key use cases.

For non-verified education, the lack of such authoritative repository must be compensated for. This is essential for example in the case where a learner has done their studies outside of the country where they are looking for study opportunities – for example when the user is an expat. In this case, the user can input the prior education in more generic terms by selecting the appropriate level and domain of the education. This information is then used in a manner roughly similar to the verified information.

The verified education scenario also offers an additional level of customisation for the user. Whereas fetching the learner's prior education data is an automated process, the way in which this information is used for the recommendations can be customised by the learner. The learner is

offered an option to mark certain units either “liked” or “disliked” by a simple thumbs up / thumbs down voting mechanism. This provides the learner an opportunity to adjust the profile and enrich the prior education information by for example emphasising topics they would like to pursue more deeply in the future or by marking themes they would like to steer away from.

Education

Finnish degrees

We have identified your following education:

Stadin ammattiopisto, Lehtikuusentien toimipaikka

Ympäristöhoitaja: Luonto- ja ympäristöalan perustutkinto
Graduated, 2016

Tell us how you liked the topics of your education, and we can give you better recommendations for learning opportunities. You can also leave the selection empty.

Kestävällä tavalla toimiminen	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Ympäristön hoitaminen	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Uusiutuvien energialähteiden hyödyntäminen	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Ulkoilureittien rakentaminen ja hoitaminen	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Kulttuuriympäristöjen kunnostaminen ja hoitaminen	<input type="checkbox"/>	<input checked="" type="checkbox"/>

Picture 1. Prior education

Module 2: Current interests

The second logical module – gathering user’s current interests – complements the education information. The aim of this prototype module is to enable the learner to express their aspirations more fully.

The interests are gathered by first presenting the user a wide range of potential fields of interest. From this pool of topics, the user can select the relevant ones. For each selected top-level field of interest, more relating topics are displayed. This allows the user to elaborate on their interests. As a result, an interest profile is formed that describes learner’s motivational topics that may have their basis on e.g. the non-formal activities such as hobbies, in addition to formal education.

Fields of interest

Next we will look at your fields of interest.

Choose 3 – 10 topics you'd like to work with.

ART AND CULTURE ✓

Linguistics ✓ Literature ✓ Music ✓ Art in general ✓ Art periods and genres ✓

Art History ✓ Architecture ✓ Visual Arts ✓ Painting ✓ Graphic Arts ✓ Sculpture ✓

Interior decoration ✓ Theatre ✓ Dance ✓ Photography ✓ Film ✓ Printing ✓

Museums ✓ Exhibitions ✓ Textile Industry ✓ Clothing Industry ✓ Footwear Industry ✓

Leather Industry ✓ Handicrafts ✓ Applied Arts ✓ Ceramics Industry ✓ Glass Industry ✓

NATURE AND ENVIRONMENT ✓

Botany ✓ Zoology ✓ Nature Protection ✓ Environmental Conservation ✓ Environment ✓

Picture 2. User's current and future interests

Module 3: Learning Opportunity Recommendations (Analytics Prototype)

The final module – the recommended education opportunities – ties together the previous two modules and completes the learner profile by offering the user meaningful analytics and insight based on their profile data. This prototype module demonstrates a concrete use case for the learner profile: by filling in the information, the user is offered recommendations for the next step on their learner path.

The recommendations use semantic matching for finding opportunities that could offer the learner new ways to combine their competences with their interests. The most suitable matches are presented for the user. The user can then examine the matches further, and also do some additional filtering such as restricting the results to a particular province.

Recommended study places

Showing the most suitable study places from your selected area based on your studies and selected fields of interests. You can restrict the area further, if you want.

Location: ▼

Tanssija

Tanssialan perustutkinto: Tanssialan perustutkinto

5 learning opportunities

Metsuri-metsäpalvelujen tuottaja

Metsätalouden osaamisala: Metsäalan perustutkinto

19 learning opportunities

Muusikko

Musiikin osaamisala: Musiikkialan perustutkinto

18 learning opportunities

Picture 3. Recommended study places

Testing and Developing Additional Analytics Features

Background

The following documentation details the joint development of analytics and AI functionalities undertaken by the University of Oulu and EDUFI. It includes a detailed description of the work and choices regarding the development of additional analytics services and collaboration with the University of Oulu. After the readjustment of financing in early 2019, a NLP (natural language processing) expert from the Faculty of Information Technology and Electrical Engineering University of Oulu was introduced to the project to add resources to the technical development of the services. As the technical developmental part of the project mainly is done by development company Reaktor, collaboration on development tasks was necessary.

At the stage of starting work, the development process was focused on the possibilities of creating a recommendation system, which would utilize user data from services and sources, namely:

1. ESCO - a multi-lingual classification of European skills/competences, qualifications, and occupations and is part of the Europe 2020 strategy. Access to ESCO is primarily in three ways: (a) as web-based API, (b) as local API, (c) as raw files downloadable from the Esco website. ESCO is divided in three separate pillars: skills/competences, occupations, and qualifications (under development). All skills, occupations, and qualifications are described as individual "concepts" that comprise of a unique identifier, description text, labels, and links to other concepts from the same or different pillar.
2. KOSKI - Finnish national register comprising educational data on all Finnish students as well as exchange students in Finland and vocational student data from the year 2018. Data in KOSKI includes student identifying data, study place and courses, final grades.
3. eRequirements –contains descriptions of degree unit parts in Finnish. The attributes obtained from the eRequirements website are: (a) degree unit number, (b) degree unit name, (c) degree unit description, (d) professional requirements, (e) ways of demonstrating skills. Optionally, information on evaluation of degree units is also available in eRequirements.

Utilization of data from KOSKI (The Finnish national register comprising vocational student data from the year 2018), eRequirements (Finnish national requirements descriptions for vocational education), ESCO (European skills competences and occupations hierarchy) and user's personal interests was to lead to the creation of study recommendations for the user on what would be an interesting type of vocational education to study. This recommendation should be based on the match of the data sets from users' interest and previous education and vocational education descriptions.

During the face-to-face meeting with EDUFI and Reaktor in Helsinki (27.3.2019), several alternative tasks for University of Oulu were discussed. These were:

1. Mapping verified competences to skills. This includes mapping competences available from KOSKI and eRequirements services to ESCO competence and skills hierarchy. Semantic matching of ESCO tags and /or their descriptions to the descriptions of completed study units.
2. Summarizing (and /or visualizing) verified competences. This task means condensation or summarization of user's competences using neural network /deep learning. The output of this would help users to realize their competences and capabilities.
3. Finding overlapping courses that can be compensated. Idea was to show users what parts of their degree could be compensated using credits from previous studies. This would require semantic matching of previous vocational degree studies to courses and units of the future degree.
4. Providing organizational structure of interest using ESCO. This means summarization or condensation of ESCO competences and skills to around 100 tags what later could be used as a basis for gathering information about users' interests.

As a result of the discussions, it was decided that the most important and urgent task for University of Oulu is to provide the organizational structure for interest (task nr.4).

Analytics test nr.1: Condensation of the ESCO Competencies to be Used as an Interest

The aim of this task was to ultimately come up with the "list" of interests to be shown to the user so that the recommendation algorithm (to be developed by Reaktor) takes interests and KOSKI profile of the user to suggest possible vocational courses. AI algorithm "crawls" ESCO and comes up with "meta" phrases that summarise what ESCO skills/occupations/qualifications are "about". For example, a job "auto mechanic" and "aircraft maintenance" are both about "working with tools".

Results

This was a challenging task, because ESCO does not include full hierarchy of competences and skills and the hierarchy that exist in ESCO cannot be used as a basis of interest from the point of view of educational psychology. For this reason, condensation of ESCO competences did not produce satisfactory outcomes. For this task to be competed outside ontology or other structuring for interest phenomena must be used together with ESCO competencies as a framework for organising and condensing competencies into relatively small number of labels.

Action Taken

Because of the limited time resources and absence of research and development regarding this particular question, other similar career guidance information databases were analysed, to understand how interest and competencies could be used together to recommend education. O*NET system widely used in US and created by the US department of labour was selected as an example of good practice. After analysing O*NET structure and methods used in their system it was concluded that this kind of otology linking interest with occupations and competences was created by human

experts manually assigning labels to job description (O*NET SOC, Judgement Method augmented with Empirical Method). O*NET's method is basing their work on Strong's Interest Inventory and its fusion into Holland Theory of vocational personality (typology), which is widely used in career guidance (Rounds, Smith, Hubert, Lewis & Rivkin, 1999 July; McCloy, Campbel, Oswald, Lewis & Rivkin, 1999). Unfortunately, ESCO and O*NET cannot be directly matched since many entities (such as jobs, qualifications, occupations) are missing from both sides to be a match.

Suggestions

For this reason, following options were presented for consideration of further possible steps:

1. Show an "interest list" derived from the standard questionnaire from Holland theory (about 30 questions, can be 30 phrases). This could give us straight and easy RIASEC code (Holland Code) that can be used in the recommendation algorithm. The drawbacks of this approach are the need to come up with the questions or statements what would be scientifically sound, would have internal validity and reliability.
2. We show the "interest list" derived from general Strong's Interest Inventory (Nauta, 2010). This would require purchasing actual inventory and would also involve the end user in answering 30 to 60 questions related to their interest.

Decisions and further development

After collaborative meetings discussing these possibilities (1-5.4.2019) it was decided that none of suggested options is suitable for the project in the current state, as they all require extensive research and collaboration work. Small local ontology *finto.fi* was decided to use instead as an outside framework to structure interest.

Analytics test nr.2: Matching eRequirements Degree Unit Descriptions to Two ESCO Pillars

During the associated partners meeting in Levi (9.4.2019) second task was suggested for University of Oulu by EDUFI and Reaktor. The goal of the suggested task is mapping existing competences to the ESCO. Specifically, this includes mapping competencies available from KOSKI and eRequirements services to ESCO competencies and skills hierarchy. This requires semantic matching of ESCO tags and /or their descriptions to the descriptions of completed study units. While main content of ESCO is in English, only the labels (given as `preferredLabel`, and `alternativeLabel`) are in non-English country-specific languages.

The task was to find mapping between eRequirements degree unit descriptions to ESCO concepts by matching descriptions of the respective datasets. Research done in preparation to this task, led to the results of other similar projects from Czech, France, Spain, and Netherlands trying to match their national education system to ESCO (European Commission, 2016). These efforts were concluded stating, that it requires "human experts" from the education sector more than 60 days to crudely map the country's system and ESCO, because of having significant differences (structure, interpretation, content) between the corresponding taxonomies.

Attempts were made to use offline ESCO Finnish version using classical NLP methods (stemming, Finnish wordnet, Turku's NLP resources) to match eRequirements descriptions to ESCO descriptions

These attempts did not yield any good matches (attached is one example of several attempts). Using Finnish wordnet harmed the performance because of semantic diversion.

Encountered problems

1. ESCO claims to have information available in 23 languages, however it does not have actual Finnish version text, just labels of the text translated to Finnish. The main textual descriptions of competences and occupations are always provided in English. Only labels (one or two words) are in Finnish. This limits the amount of text possible to use in semantic matching and does not produce satisfactory results.
2. The search in the downloaded API version of ESCO does not work for Finnish language (see point 4), although it works for many other European languages.
3. There are gross differences between the interpretation, discourse style, and intended purpose of eRequirements descriptions and ESCO concepts. For this reason, there is no possibility to search relationship between categories (competencies and occupations). The search is only possible inside the categories separately.
4. Finnish is a morphologically rich language with many inflections which makes it difficult to compare the words directly from one narration to the other.

Approaches to solve these problems

1. Direct matching words from eRequirements descriptions to ESCO:
 - Both eRequirements description and ESCO labels are considered as two bags of words
 - Similarity is essentially the intersection between these two sets of words (although, word order is respected).
2. Stemming and matching:
 - Both eRequirements description and ESCO labels are processed to find the "stems" of the words (the least common prefix among all variations of a given word).
 - These two sets of stems are then evaluated to find similarity between them.
3. Lemmatizing and matching:
 - Both eRequirements description and ESCO labels are processed to find the "lemmas" of the words (the root form of the word among all variations of a given word).
 - These two sets of lemmas are then evaluated to find similarity between them.
4. Distributional semantic similarity-based matching:
 - Each word from eRequirements description and ESCO labels was mapped to a vector (word embedding vector) using word2vec model pre-trained on 4.5 billion words by Turku NLP group.
 - These vectors are then compared using Wasserstein's similarity criterion.

5. Translating eRequirements descriptions to English and then matching:
 - All descriptions of eRequirements were translated into English (from Finnish) by a translation API.
 - These translated descriptions are then matched with ESCO English results descriptions using bags-of-words model.

Results

Units from eRequirements have been matched to ESCO competences or/and occupations using variety of methods. Results of these approaches have been evaluated by the CSC, EDUFI and Reaktor for further planning and possibility to incorporate them into the prototype before the end of the project.

Summary of the results

One of the main goals of WP3 was to show a user's competence profile. We gathered information from the Koski register and we tried to enrich that information with a another way showing competences. Those ideas and ways are described earlier in this document. The main idea was to map study units to the ESCO classifications. This research work was done by the Oulu university in collaboration with EDUFI and Reaktor. However mapping was proofed to be too time consuming and complex in the projects time frame.

EDUFI also examined two other options of enrichening the competence profile with ESCO mapping. These alternatives which were provided by other vendors than Oulu university or Reaktor had other problems and challenges and that's why those options had to ruled out. Especially the way of mapping the competences to ESCOs was unfortunately not fruitful enough for the project, and only resulted in ESCO suggestions that were unhelpful for the users.

Unfortunately the work of breaking down competences in the University of Oulu was thus unsuccessful byt the time of this submission. The quality of the output competences was too low to be useful as a feature for the POC and the integration was thus not carried out.

References

1. McCloy, R., Campbel, J., Oswald, L. F., Lewis, P. & Rivkin, D. (1999). Linking Client Assessment Profiles to O*NET™ Occupational Profiles. National Center for O*NET Development Employment Security Commission. Raleigh, North Carolina
2. Nauta, M. M. (2010). The Development, Evolution, and Status of Holland's Theory of Vocational Personalities: Reflections and Future Directions for Counseling Psychology. *Journal of Counseling Psychology*, 57(1), 11–22.

3. Rounds, J. Smith, T., Hubert, L., Lewis, P., & Rivkin, D. (1999 July). Development of Occupational Interest Profiles for O*NET. National Center for O*NET Development Employment Security Commission. Raleigh, North Carolina.
4. European Commission (2016). ESCO Mapping Pilot – Final Report. Employment, Social Affairs & Inclusion