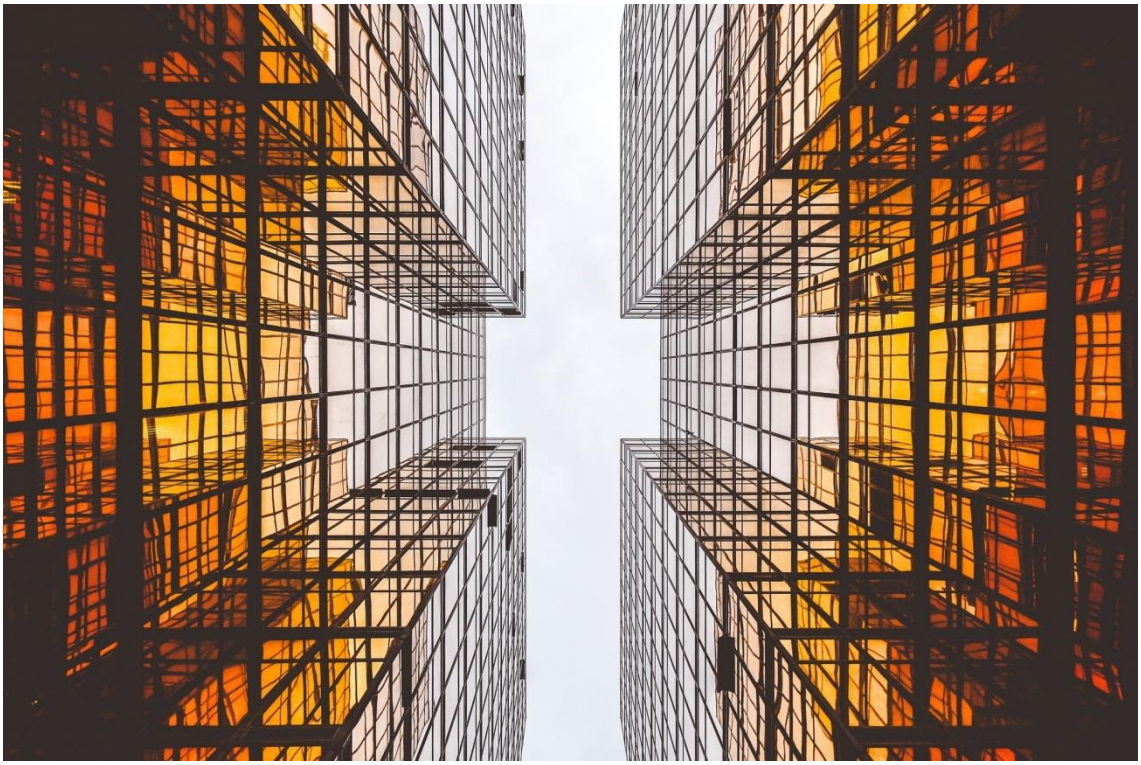

D9 REPORT OF TECHNOLOGICAL REQUIREMENTS OF THE SYSTEM

Report	WP3 Methodological design of the tool for detection of competencies and training needs
Version	2.0
Status	Final
Date of issue	
Dissemination level	Public
Authors	Anselmo Navarro, Ángel Esteban (FORTEC), David Abadía, Paula Peña (ITAINNOVA),
Contributors	
Project Co-ordinator	Anselmo Navarro Martínez (FORTEC)
Project website	www.e-detecta.eu

¶

Co-funded by the Erasmus+ Programme of the European Union

This·project·has·been·funded·with·support·from·the·European·Commission.·This·publication reflects·the·views·only·of·the·author,·and·the·Commission·cannot·be·held·responsible·for·any·use·which·may·be·made·of·the·information·contained·therein.·¶

Project·Number:59184-EPP-1-2017-1-ES-EPPKA2-SSA-N¶

Disclaimer:

This publication has been produced in the frame of the DETECTA Project. The project has been funded with support from the European Commission, ERASMUS+ Programme. The information contained in this publication reflects only the author's views and the Commission cannot be held responsible for any use which may be made of the information contained therein. Sole responsibility lies with the authors.

TABLE OF CONTENTS

1 EXECUTIVE SUMMARY

Within this Deliverable the technological requirements of the DETECTA system are presented. These requirements are used as a base for the posterior developments concerning the DETECTA tool by aiming at the technologies to be used, definition of the user types to cope with the tool, which requirements are to be consider from a user perspective, accommodation requirements for the tool, description of the test to verify the right behavior of the tool as well as the criteria of acceptance.

2 INTRODUCTION

The goal of the present Deliverable D9 REPORT OF TECHNOLOGICAL REQUIREMENTS OF THE SYSTEM is to identify the technical requirements concerning the development of the DETECTA system. This goal covers different scopes which are presented below:

- Technologies considered during the development of the DETECTA project: Keeping in mind that the main goal of DETECTA platform is to recommend a training itinerary to occupation seekers of the Construction sector who want to achieve a specific occupation position, involved technologies to achieve this purpose are presented, concerning among others crawler models (to retrieve information from external data), involved data sources (based on functionalities to achieve) concerning job offers as well as training offers, ontology (to be used to classify occupations, competences and qualifications of the workers, by considering ESCO ontology as a base to build on the DETECTA ontology), data storage (used to store the information extracted from external data sources within the DETECTA platform) as well as the hosting technologies (to assure redundancy, availability and access to multiple devices concurrently).
- Definition of user types: presenting end users who will exchange information with the system, i.e. the actors of the system in order to clarify the relation among themselves and the activities involved in the use cases.
- User requirements: a set of user requirements (functional and non-functional) defined for the DETECTA system is presented, where functional requirements deal with functionality that is visible and important to users that the system has to deliver to satisfy the business objectives that the system is designed to fulfill and non-functional requirements deal with the characteristics that the system must meet to provide a sufficient level of reliability, usability, maintainability, availability, etc., which are important for proper operation of the system.
- Criteria of acceptance, where acceptance criteria corresponding to the defined functional requirements (criteria are conditions that a software product must satisfy to be accepted by a user, a customer or other system).
- Accommodation requirements defined for the proposed infrastructure of the DETECTA system.
- Description of tests to develop in order to verify the right working of the tool, where different test phases that will be carried out in the DETECTA system through the V-model are described.

Based on the purpose of this document (which is to identify the technical requirements of the DETECTA system) the target groups for this document are related to:

- On one hand developers of the DETECTA system, who participate during the development of the DETECTA system, so that the technical requirements for the development are compiled to cope with the functionality pf the proposed system.
- On the other hand IT administrators of the DETECTA system, who are IT responsible persons for the deployment and the maintenance of the DETECTA system so that they are aware of the involved technologies.

This report (already sent) has been revised and updated by subsequent investigations accordingly, this being the final version.

3 CONSIDERED TECHNOLOGIES

In this section, considered technologies for the development of the DETECTA system are presented. Main goal of DETECTA platform is to recommend a training itinerary to occupation seekers of the Construction sector who want to achieve a specific occupation position, by taking into account:

- User's profile (job seeker, employee or company)
- User's desired occupation
- Job and training offers in market related to desired occupation position
- Reports which enable to analyze trends
- Skills and competences related to users, companies, job and training offers

Main functional modules are presented in the Figure 1 below:

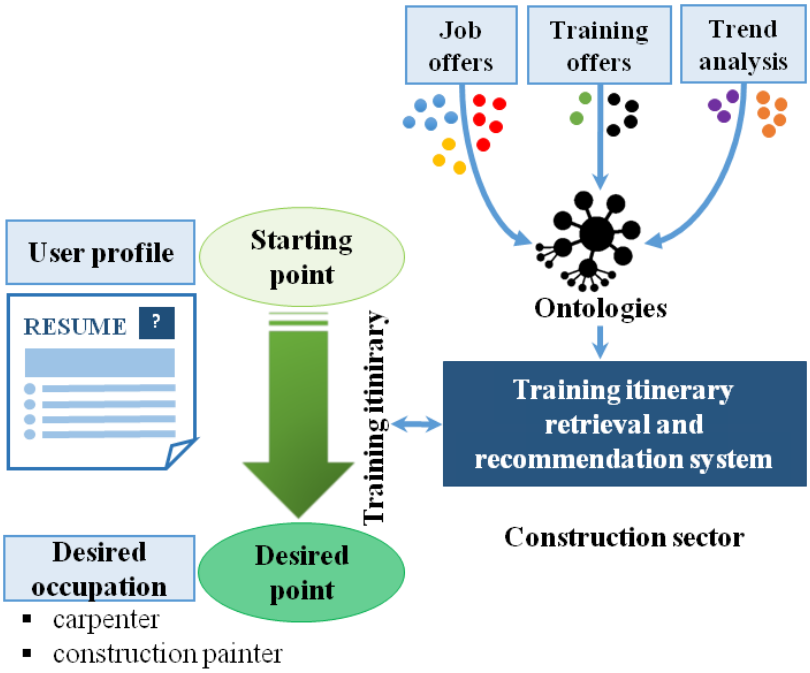

Figure 1: Recommendation process embedded in DETECTA system.

In Figure 1, it is shown the recommendation process embedded in DETECTA system, which supports a recommendation model that has as its **starting point** (or input):

- The **user's profile**, that contains information related to the job seeker, employee or company, such as personal information (e.g., name, surname, address, country, region, website and phone number), personal skills (e.g., driver license, language, experience and

occupations) and education (e.g., training centres, professional and unprofessional training courses).

- The **desired occupation** by the user (job seeker, employee or company) in the construction field (e.g. building labourer, supervisor, civil engineer, etc.).
- **Geo location** constraints (e.g. current country).
- **Language** constraints (e.g. English, French and Spanish).
- **Number of training courses** to retrieve during the search

The **output** of the recommendation system is a set of **training courses** (or training itinerary) that the current user lacks, in order to achieve the user's desired job. For each recommended training course, the user can access information related to training modules, job offers, occupations and training centres. During the recommendation process, the model is fed with external data sources related to job and training offers, as well as trends analysis. This will be possible thanks to combination of semantic intelligence technologies with the Big Data's large-volume information analysis power.

In order to carry out the development of the DETECTA system, a set of accessible and easy-to-deploy technologies will be used, such as the following:

- **Crawler model:** It will require the implementation of several crawlers that will include the following four steps:

    1. **Access:** The crawler will have a connection module to the relevant external data source (or selected web portal). The possibilities can be a web connection or through a REST API.

    2. **Acquisition:** The information will be collected and stored in a variable or in a physical location of the server. Each data source can have a different format. Hence, it will be necessary to create an acquisition algorithm for each web portal. In the case of existing data sources with similar formats (or equals), the same crawler may be used.

    3. **Transformation:** Through this module, the crawler will be able to transform the initial format of the information extracted, to the format required by the system.

    4. **Data storage:** The obtained information will be stored in a database.

    The system will have parameterized one module for each data source (or web portal) and the methods used can be the following:

- **Direct connection through a curl:** The system will have a connection between servers through URL libraries, which allows to make different requests, such as get, post, put or delete. The first and second allows requests to retrieve the data from the web page; the third is used by the system to store information; and the last, normally, is used to delete some data from the system.
  In this type of connection, the retrieved data are in unstructured formats and it is necessary to make a transformation of this initial format to a format known by the system to be developed. In these cases, normally, in the phase of accessing and obtaining information, it is necessary to apply some HTML library that allows retrieving and parsing the HTML content.
- **API-s:** It may require the development of libraries on the platform or use external libraries to make connections with certain relevant external data sources and facilitate the transfer of data between systems. In general, these libraries have documentation, they are tested and the results they return are of quality.

- **Data sources:** Among the different existing web portals, will be selected, as far as possible, those that meet the following criteria:
  - Legal issues to access data sources
  - European web portals, that cover the three languages (Spanish, French and English) and 12 countries considered.
  - Transversal public or private portals, with the same information structure.
  - Available the geo location of job and training offers.

  The different types of web portals to be considered will be the following:
  - **Job offers:** In order to extract relevant information about job offers, several web portals could be used, such as:
    - **EURES[1]:** It is job mobility web portal that has hosts more than one million job vacancies from all over Europe, almost as many CVs, and thousands of registered employers. The content of job offers can be easily obtained in a homogeneous format (JSON). In addition, it use the ESCO ontology[2], as a standardized terminology that improves the semantic interoperability and facilitates the description of occupations, skills, competences and qualifications, which are linked and interact with each other.
    - **Indeed[3]:** It is an employment-related search engine that includes job listings from thousands of websites. It has a job search API well documented, but a user account and application credentials are

[1] https://ec.europa.eu/eures/eures-searchengine
[2] https://ec.europa.eu/esco/portal/occupation?resetLanguage=true&newLanguage=en
[3] https://www.indeed.es

required. In addition, the maximum number of jobs retrieved per query is limited to 1000.

- **Empleate**[4]**:** It is an employment-related search engine for the Spanish job market. However, job offers are not limited to the national geography; there are offers in the European Union, the United States or Latin American countries. In addition, it integrates job offers from private web portals. To date, access is public through queries to the Solr database.
- **InfoJobs**[5]**:** It is a private employment portal mainly designed for the Spanish, Italian and Brazilian job market. It has a Job Search API with available documentation, but requires a user account and application credentials.
- **Monster**[6]**:** It is a global employment website that includes job offers of more than 40 countries. The Job Search API requires a user account. In addition, the maximum number of jobs retrieved per query is limited to 1000.

- **Training offers:** In order to extract relevant information about training offers, several web portals could be considered, such as:
  - **Drop'pin@EURES**[7]**:** It is a web portal where companies and organizations can promote and show youth opportunities. It is designed to help young Europeans take their first steps into the job market. Opportunities on this online platform include apprenticeships, traineeships, training offers, e-learning courses, language training, mobility support, coaching and mentoring etc.
  - **Ploteous**[8]**:** It is a web portal that contains opportunities for learning and training in Europe. The web portal is divided in different sections: learning opportunities, qualifications, guidance services, exchange programs and Erasmus plus other Grants and recognition of diplomas and qualifications.

- **Professional certificates:** The web portals that contains information related to professional certificates (or education levels) could be considered:
  - **Professional Certificate of SEPE (Public Employment Service in Spain)**[9]**:** It is official accreditation instrument for the professional qualifications of

[4] https://www.empleate.gob.es
[5] https://www.infojobs.net
[6] https://www.monster.es
[7] https://ec.europa.eu/eures/public/es/opportunities
[8] https://ec.europa.eu/ploteus
[9] https://www.sepe.es/HomeSepe/Personas/formacion/certificados-profesionalidad/familias-profesionales/edificacion-obra-civil.html

the National Catalog of Professional Qualifications in the field of job administration. These certificates accredit a set of professional competences that enable the development of an identifiable work activity in the productive system without this constituting a regulation of the professional exercise. They have official status and validity throughout the national territory and are issued by the SEPE and the competent organisms of the Autonomous Communities.

- **Europass Certificate Supplements**[10]**:** National inventories of Certificate Supplements, which are documents that describe the knowledge and skills acquired by holders of vocational training certificates.

- **Trend analysis:** In order to achieve greater representativeness, the trend analysis in the construction field could be based on interviews and social media analysis (e.g. twitter, newspaper articles, reports, etc.).

The crawler module will facilitate the automatic inspection of the content published on the web portals finally selected. The web portals related to training offers, professional certificates and job market trends are not linked to the ESCO ontology. Hence, a standardization of skills and competences should be done previously, by using **textual classification models**. In Figure 2, the ESCO ontology and its relationship with job and training offers is shown.

[10] https://europass.cedefop.europa.eu/documents/european-skills-passport/certificate-supplement/national-inventories

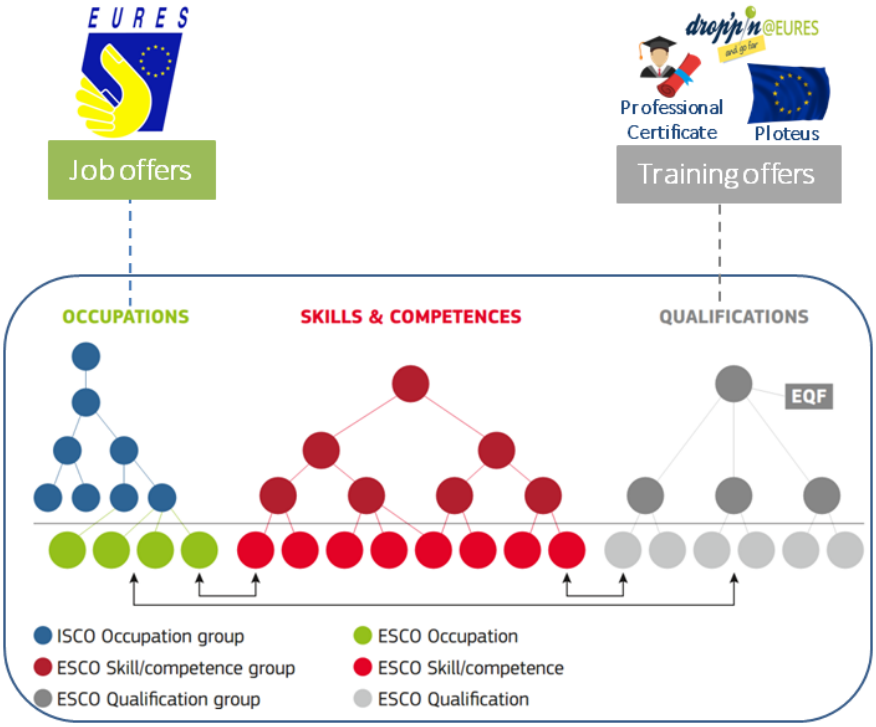

Figure 2: Three pillars of the ESCO taxonomy and its relationship with job and training offers.

- **Ontology:** An ontology is going to be used to classify occupations, competences and qualifications of the workers, ESCO Ontology is proposed to be used to define the DETECTA ontology as a basis as it is based on ESCO taxonomy that contains a multilanguage classification of European skills, competences, occupations and qualifications. Actually, the ESCO taxonomy contains[11]:

    - **2942 occupations:** It organizes the occupation concepts in ESCO. It uses hierarchical relationships between them, metadata as well as mappings to the International Standard Classification of Occupations (ISCO) in order to structure the occupations. Each occupation concept contains one preferred term and any number of non-preferred terms and hidden terms in each of the ESCO languages. Each occupation also comes with an occupational profile. The profiles contain an explanation of the occupation in the form of description, scope note and definition. Furthermore, they list the knowledge, skills and competences that experts considered relevant terminology for this occupation on a European scale.

    - **13485 skills/competences:** ESCO distinguish between the skill/competence concepts and knowledge concepts, but it does not distinguish between skills and

[11] https://ec.europa.eu/esco/portal/occupation

competences. Each ESCO occupation has a description, notes, labels, essential and optional competences and knowledge. This data is used to build the ontological base.

- 8677 qualifications: Qualifications are the formal outcome of an assessment and validation process, which is obtained when a competent body determines that an individual has achieved learning outcomes to given standards. Qualifications displayed in ESCO come from databases of national qualifications that are owned and managed by the European Member States. Member States provide this information to ESCO on a voluntary basis. It therefore depends on each Member State to ensure information on its qualifications in ESCO is available, complete, correct and up-to-date. The Commission also envisages integrating private, international and sectorial qualifications from other sources into ESCO in the near future. It is piloting this approach and discussing it with the Member States.

Different elements ensure the quality of ESCO[12]:

- Data model: Define the data model and the business rules allowing ESCO users and developers to work efficiently with the vocabulary.
- ESCO guidelines: Define guidelines for the development of ESCO and make them available to all experts involved.
- Experts' involvement: Involve experts from all over Europe and from all sectors of economic activity to develop and/or evaluate the draft occupations, knowledge, skills and competences.
- Quality assurance steps and supporting tools: Define quality assurance steps, checklists, reporting templates, and tools for the Commission and the ESCO Maintenance Committee to monitor the work done and validate the deliverables.
- Mapping to ISCO-08: Quality assure the mapping to ISCO-08 with the support of an expert suggested by the ILO.
- Completeness: Complete a gap analysis between the ESCO classification, ISCO-08 and eight national classifications to: ensure the completeness of ESCO; identify discrepancies in the level of detail between ESCO, ISCO-08 and the national classifications; enrich the ESCO classification with additional non-preferred terms.
- Avoid ambiguities in the occupation structure: Check the scope alignment between the ESCO English terms used in the ESCO occupations and ISCO-08.

[12] Handbook, E.S.C.O. European Skills, Competences, Qualifications and Occupations (2017). *European Comision. Directorate E*.

- **Terminological guidelines:** Involve linguists to develop general and language-specific guidelines for the formulation of terms in the various languages.
- **Involve professional translators:** Have translators formulate the terms for the various languages.
- **Provide the translators with access to job market experts:** Provide the translators involved in the formulation of terms with access to job market experts in a given language, to assist them in identifying the terms really used in the job market.
- **Feasibility check with the Member States:** Consult the Member States on the feasibility of mapping national classifications to ESCO.
- **Language check with the Member States:** Consult the Member States on the accuracy of the ESCO terms in its 25 language variants.

- **Data storage: A database is going to** be used to store the information extracted from external data sources within the DETECTA platform. Requirements concerning scalability, security, performance and cost-effectiveness solutions are taking into account, so an open source database is proposed to be used. Among the existing open sources databases Solr is chosen as it is an open source NoSQL database system, which stores data structures in documents with a dynamic schema, making easier and faster the integration of data. The version to use of Solr must support texts and application of related Natural Language Processing techniques in multiple languages (e.g. English, Spanish and French) as well as provide scalable indexing and advanced analysis/tokenization capabilities.

- **Hosting:** The DETECTA platform will be deployed in such a way to assure redundancy, availability and access to multiple devices concurrently. Proposed deployment will be held on a cloud based architecture, e.g. Google, Amazon or Azure, based on Linux. DETECTA system must then be installed in a machine (integrated within the cloud platform) with Linux operating system due to its robustness, scalability and technical knowledge for system administrators.
  In addition, the contracted server must have the following features:
  - a Linux machine with 20 cores,
  - 64 gigabytes of physical memory,
  - 20 terabytes of hard disk,
  - a transfer rate of 1 Gbps (one gigabit per second),
  - connection protocol open via SSH (protocol remote administration that allows users to control and modify their remote servers over the Internet).

  If it is necessary, because we increase the number of countries and users, the system will scale to 4 or more machines working in parallel, thanks to the scalability of the proposed solution.

4 DEFINITION OF USER TYPES

In this section, end users who will exchange information with the system are presented. The actors of the system clarify the relation among themselves and the activities involved in the use cases, in order to provide the definition and representation of the functional requirements through the use cases of the system.

The diagram of actors of the DETECTA system is shown in Figure 3.

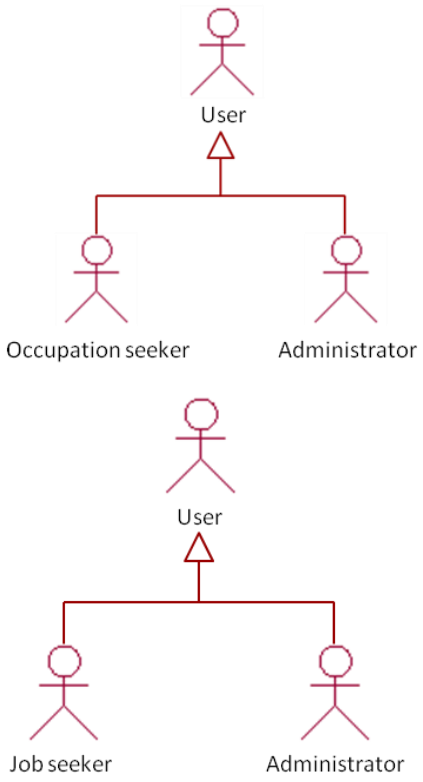

Figure 3: Diagram of actors of the DETECTA system.

In order to facilitate the understanding of the role that users play in the interaction with the system, they are described in the following tables:

Name	AC01. Occupation seeker
Description	It represents any person (job seeker or employee) or company of the Construction sector who requires recommendations of training itineraries to achieve an occupation position.

Name	AC02. Administrator
Description	It represents the person in charge of configuring modules to access URLs of the web portals, solving end-user issues and monitoring the overall performance of the system (e.g. performance of data downloads, use of database storage, availability of computer usage, etc.), in order to detect possible anomalies. In addition, the administrator can visualize information related to training and occupation needs of occupation seekers in the Construction sector.

5 USER REQUIREMENTS

In this section, a set of user requirements (functional and non-functional) defined for the DETECTA system is presented.

On the one hand, the functional requirements deal with functionality that is visible and important to users that the system has to deliver to satisfy the business objectives that the system is designed to fulfill. Based on functional user requirements, the DETECTA system should be able to provide a user following capabilities:

FR1. Register of user
FR2. Recommend training itinerary
FR3. Manage system
FR4. Statistical reports

On the other hand, the non-functional requirements deal with the characteristics that the system must meet to provide a sufficient level of reliability, usability, maintainability, availability, etc., which are important for proper operation of the system. They are general characteristics of the system and may not be necessarily directly observable by the user.

The non-functional user requirements of the DETECTA system are mentioned below:

NFR1. **Usability:** The design of the web application should be simple and friendly to speed up the connection time to it, as well as facilitate user navigation through the application.
NFR2. **Performance:** The response time of recommendation to the occupation seeker about training itineraries must be quick. Hence, the recommendation model must be efficient.
NFR3. **Portability:** The DETECTA system must be multiplatform.

NFR4. **Security:** The system must guarantee the privacy of users' data. However, it is necessary to know the user's geolocation in order to take it into account during the recommendation process.

NFR5. **Help manual:** The system must have an administration and job seeker manual in order to provide guidance to users regarding system options, using explanatory texts that indicate the action of these.

NFR6. **Software:** The access to the application will be through of different devices (e.g. computers, tablets, smartphones, etc.).

NFR7. **Hardware**: The system will be deployed in an Amazon cloud infrastructure on a Linux machine with 20 cores, 64 GB of memory and 20 Terabytes of HD.

NFR8. **Language:** The application must support English, Spanish and French languages.

6 CRITERIA OF ACCEPTANCE

In this section, acceptance criteria corresponding to the defined functional requirements (see Section 4) are described.

The criteria are conditions that a software product must satisfy to be accepted by a user, a customer or other system. They are unique for each user story and define the feature behavior from the end-user's perspective. Well-written acceptance criteria help avoid unexpected results in the end of a development stage and ensure that all stakeholders and users are satisfied with what they get.

There are different acceptance criteria types and structures, such as:

- **Scenario-oriented (Given/When/Then)**: It has the following structure:
  - *Given* some precondition
  - *When* I do some action
  - *Then* I expect some result

  and provides a consistent structure that helps testers define when to begin and end testing a particular feature. It also reduces the time spent on writing test cases as the behavior of the system is described upfront.

  Each acceptance criteria written in this format has the following statements:

  1. **Scenario**: the name for the behavior that will be described
  2. **Given:** the beginning state of the scenario
  3. **When:** specific action that the user makes
  4. **Then:** the outcome of the action in "When"
  5. **And:** used to continue any of three previous statements

When combined these statements cover all actions that a user takes to complete a task and experience the outcome.

- **Rule-oriented (checklist)**: It entails that there is a set of rules that describe the behavior of a system. Based on these rules, specific scenarios can be drawn. Usually, criteria composed using this form look like a simple bullet list.

The scenario-oriented acceptance criteria for each functional requirement of the DETECTA system are mentioned below:

Functional requirement	Scenario	Criteria of acceptance
FR1.1	The occupation seeker makes a valid registration.	Given that the occupation seeker registration is valid, when: • the system verifies that the current occupation seeker does not exist, • the user fills a form related to his/her profile (e.g. basic personal data, cv information, etc.), and clicks the "save registration" button, then register user is saved.
FR1.2	The occupation seeker updates her/his profile.	Given that the occupation seeker profile is updated, when: • the occupation seeker incorporates new information in his/her profile, • the system automatically re-trains the recommendation model, considering the information of the user's current profile, and • the system stores the new recommendation model for later use, then the user's profile is updated.
FR2.1	The occupation seeker explicitly defines a valid recommendation search.	Given the recommendation search is valid, when the occupation seeker introduces: • the desired occupation type, • context restrictions and

		• number of training courses, and clicks the “search” button, then the recommendation search is performed.
FR3.1	The administrator explicitly configures valid URLs.	Given the URLs are valid, when the administrator introduces URLs of web portals with information related to: • Job offers • Training offers • Professional certificates and clicks on the “save” button. Then the system verifies that URLs of web portals are available for downloading information and initiates the process of extracting and storing of information in the database.
FR3.2	The administrator integrates new data sources.	Given new data sources integrated into the system, when the administrator: • introduces new URLs of web portals, • updates the recommendation algorithm to the new input data sources, • stores in the database new URLs (new web portals), and • incorporates in the database the necessary fields to store the information extracted from web portals, then the new data sources are integrated in the system.
FR4	The administrator or occupation seeker accesses to statistical reports.	Given data of the different sources integrated into the system, and the queries of the job seekers, when the administrator or occupation seeker: • accesses to general reports of data sources,

		• accesses to specific reports by language, then the data are integrated in the reporting.

The defined acceptance criteria are simple, approachable and allow solving multiple problems at the same time. They will facilitate the following:

- document customer expectations
- provide an end-user perspective
- clarify requirements
- prevent ambiguity
- help quality assurance verify if the development goals were met

7 ACCOMMODATION REQUIREMENTS

In this section, the accommodation requirements defined for the proposed infrastructure of the DETECTA system are presented. Regarding to storage and computing servers (hardware), two environments will be deployed:

- **Development environment:** It will be deployed in "Super Computing Center of Aragon" (CESAR), and will be used during the development of the system.

- **Production environment:** It will be deployed in an infrastructure compatible with the national security scheme of a cloud service provider, as is the case of the Amazon cloud services. In this way, quality will be guaranteed and the maintenance infrastructure will be reduced.

In the development environment, Moriarty® will be deployed as the software development tool. It is owned by ITAINNOVA, which is the result of more than 10 years of research in semantic analysis and artificial intelligence. Moriarty® allows to develop applications that learn, understand patterns, and help decision-making, by using several information sources. Its architecture integrates six modules:

- **social:** information analysis of social networks

- **search**: intelligent semantic search engine
- **suggesting**: automatic recommender
- **profiling:** user profiler by category
- **insights analytics:** analysis and text comprehension

These modules will allow to develop the requirements of DETECTA in a very short time.

Some of the advantages of using Moriarty® are the following:

- Drastic reduction in the development of the system. This framework integrates developed artificial intelligence and big data services for the analysis of textual information and data. It allows to reuse these services and visualize the first results without having to develop the whole solution from scratch.
- It allows developing real-time systems, and large amounts of data.
- The scalability of the system is wide through the use of Big Data architecture. It is also robust and multi-user.
- The maintenance of the system is simple (one single system, in multiple clients).
- High integration with other Big Data solutions from third parties.
- In the case that the information analysis requirements change, the models are changed transparently for the end users, who can continue using the application at all times.

It allows the rapid development of prototypes based on results obtained from initial phases of prototypes, which are adapted to the defined requirements.

8 DESCRIPTION OF TESTS TO DEVELOP IN ORDER TO VERIFY THE RIGHT WORKING OF THE TOOL

In this section, the different test phases that will be carried out in the DETECTA system through the V-model are described. It is a linear development methodology used during the software development life cycle. This methodology specifies various linear stages that should occur across the life cycle, one at a time, until the project is complete (see Figure 4). A component unique to the V-model is that during each design stage, the corresponding tests are also designed to be implemented later during the testing stages.

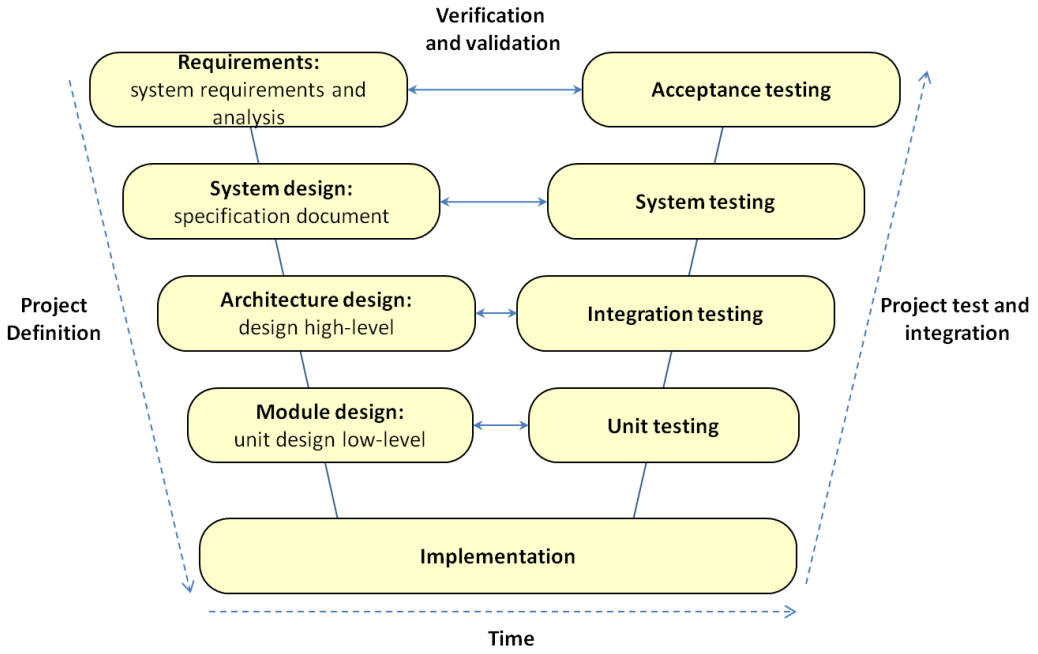

Figure 4: V-model life cycle.

The V-model includes stages such as the following:

- **Requirements:** In the initial phase, system requirements and analysis are performed to determine the feature set and needs of users. In this stage, acceptance tests are designed.
- **System design:** In this phase, it is generated a specification document that will outline technical components (e.g. data layers, business logic, etc.), by using the feedback and user requirement documents created during the requirements phase. System tests are also designed during this stage for later use.
- **Architecture design:** In this phase, specifications are drawn up that detail how the application will link up all its various components, either internally or via outside integrations. Often this referred to as high-level design. Integration tests are also developed during this time.
- **Module design:** This phase consists of all the low-level design for the system, including detailed specifications for how all functional, coded business logic will be implemented (e.g. models, components, interfaces, etc.). Unit tests should also be created during the module design stage.
- **Implementation:** In this phase, the implementation takes place. This period should allot for as much time as is necessary to convert all previously generated design and specification docs into a coded, functional system. This stage should be fully complete once the testing phases begin.
- **Unit testing:** In this phase, the process moves back up the far side of the V-model with inverse testing, starting with the unit tests developed during the module design

phase. This phase should eliminate the vast majority of potential bugs and issues, and thus will be the lengthiest testing phase of the project. In general, this stage will allow verifying that the internal business logic of the application works correctly.

- **Integration testing:** In this phase, the testing defined during the architecture design phase is executed, ensuring that the system functions across all components and third-party integrations.
- **System testing:** In this phase, the tests created during the system design are executed, largely focusing on performance and regression testing.
- **Acceptance testing:** In this phase, the acceptance tests created during the initial requirements phase are executed, and should ensure that the system is functional in a live environment with actual data, ready for deployment.

The V-model offers a set of **advantages** over other systems development models, among which the following are mentioned:

- The users of the V-model participate in the development and maintenance of the model.
- The V-model provides concrete assistance on how to implement an activity and its work steps, defining explicitly the events needed to complete a work step: each activity schema contains instructions, recommendations and detailed explanations of the activity.
- The V-model works well for projects where requirements are well defined.
- The V-model is an extension the water model, but the time concern in comparison with the waterfall model is low. In addition, the V-model has higher chance of success over the waterfall model due to the development of test plans early on during the life cycle.

During DETECTA system development lifecycle, the V-model will be carried out. The left side of the model represents the software development life cycle. It is the sequence of activities (e.g. requirement analysis, system design, architecture design, design of modules and implementation) that the developers will carry out in order to design and develop high-quality software. In addition, this side also test cases are created.

On the other hand, the right side of the model is software test life cycle. It consists of a series of activities (e.g. acceptance testing, system testing, integration testing and unit testing) carried out by testers methodologically to test the system. During these activities the test cases will be executed by developers. In addition, tests will be carried out by expert users (consortium members) and end users (job seekers), in order to validate the system in different control environments.

9 CONCLUSION

In this report, the technological requirements of the DETECTA system were presented. Goal of DETECTA is to achieve a recommendation system capable of suggesting training courses to occupation seekers (job seekers, employees and companies), who want to achieve a specific occupation position in the Construction sector. Firstly, considered technologies for the development of the system were proposed. Secondly, end users who will interact with DETECTA system were described. Thirdly, a set of user requirements (functional and non-functional) were defined. In order to validate the defined user requirements, a set of scenario-oriented acceptance criteria has been proposed. Then, accommodation requirements defined for the proposed DETECTA infrastructure were presented. Finally, the V-methodology to performance during the DETECTA development life cycle was explained.