Big Data for Labour Market Intelligence

Capacity development programme 2022

Module 1: Technical training
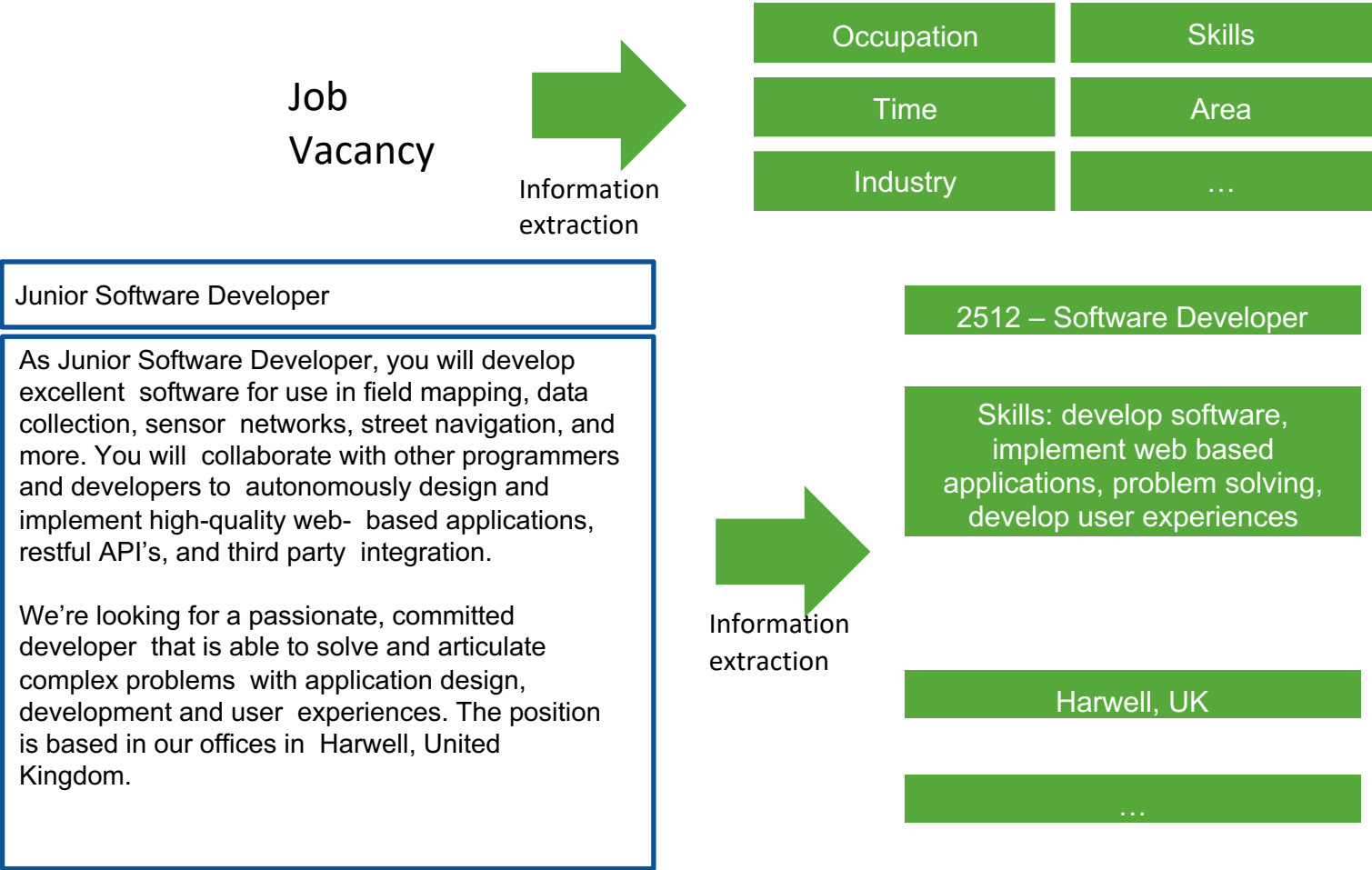
Session 6

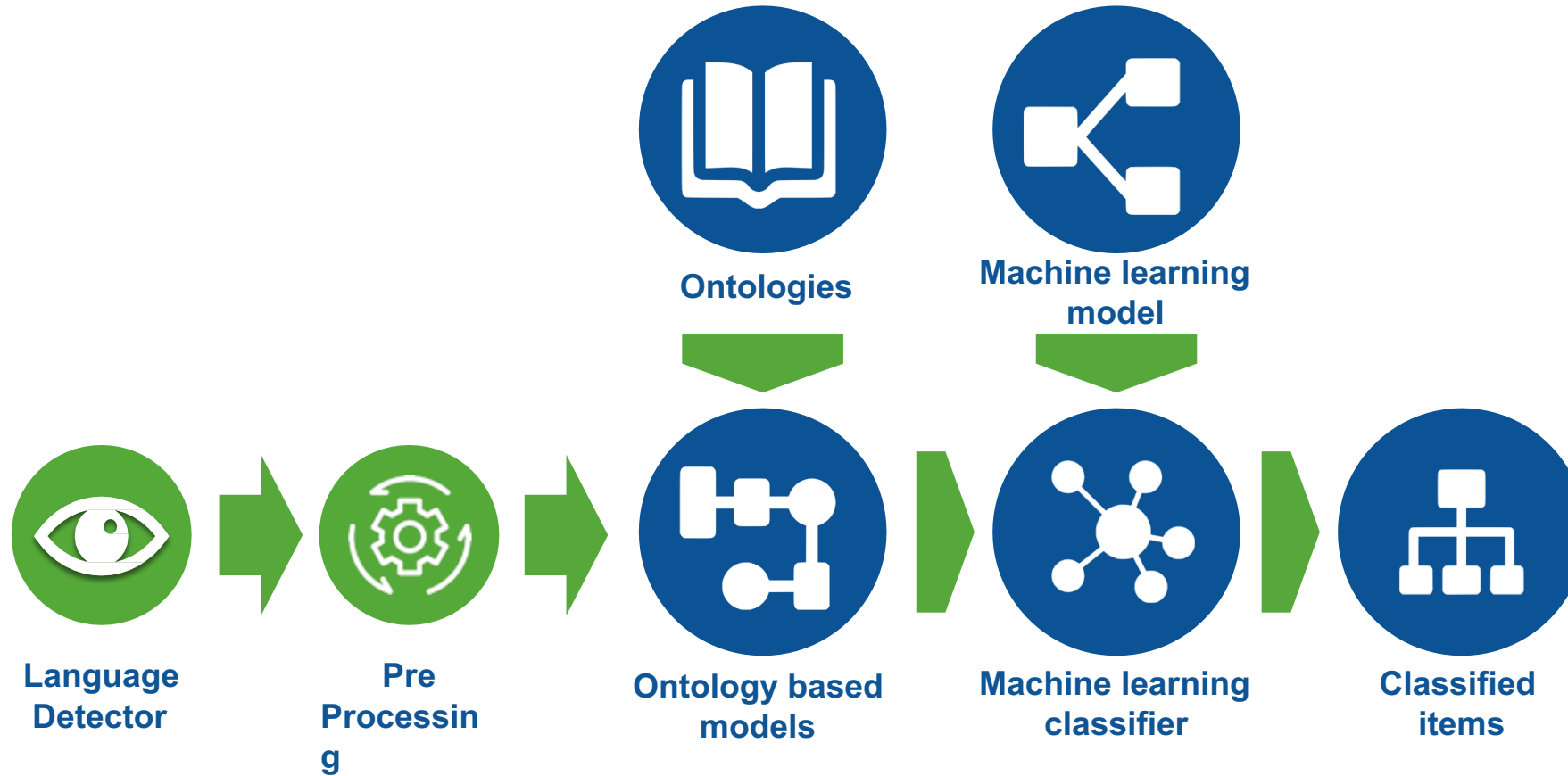**Data classification, data retrieval, practical exercise and coding hands-on**

Speaker: Simone Perego

09/11/2022

Data Classification

- ### Goal:
  - **Extract and structure information from data, to be provided to the presentation layer**
- ### Challenges:
  - **Handle massive amount of heterogeneous data written in different languages**
- ### Approach:
  - **Develop an adaptable framework, language dependent, tailored on different information features. Some relevant challenges:**
    - **Occupation feature classification: combined methods such as Machine Learning, Topic Modeling and Unsupervised Learning**
    - **Skill feature classification: another different combined methods, such as Text Analysis with corpus based or Knowledge based similarity**
- ### Features:
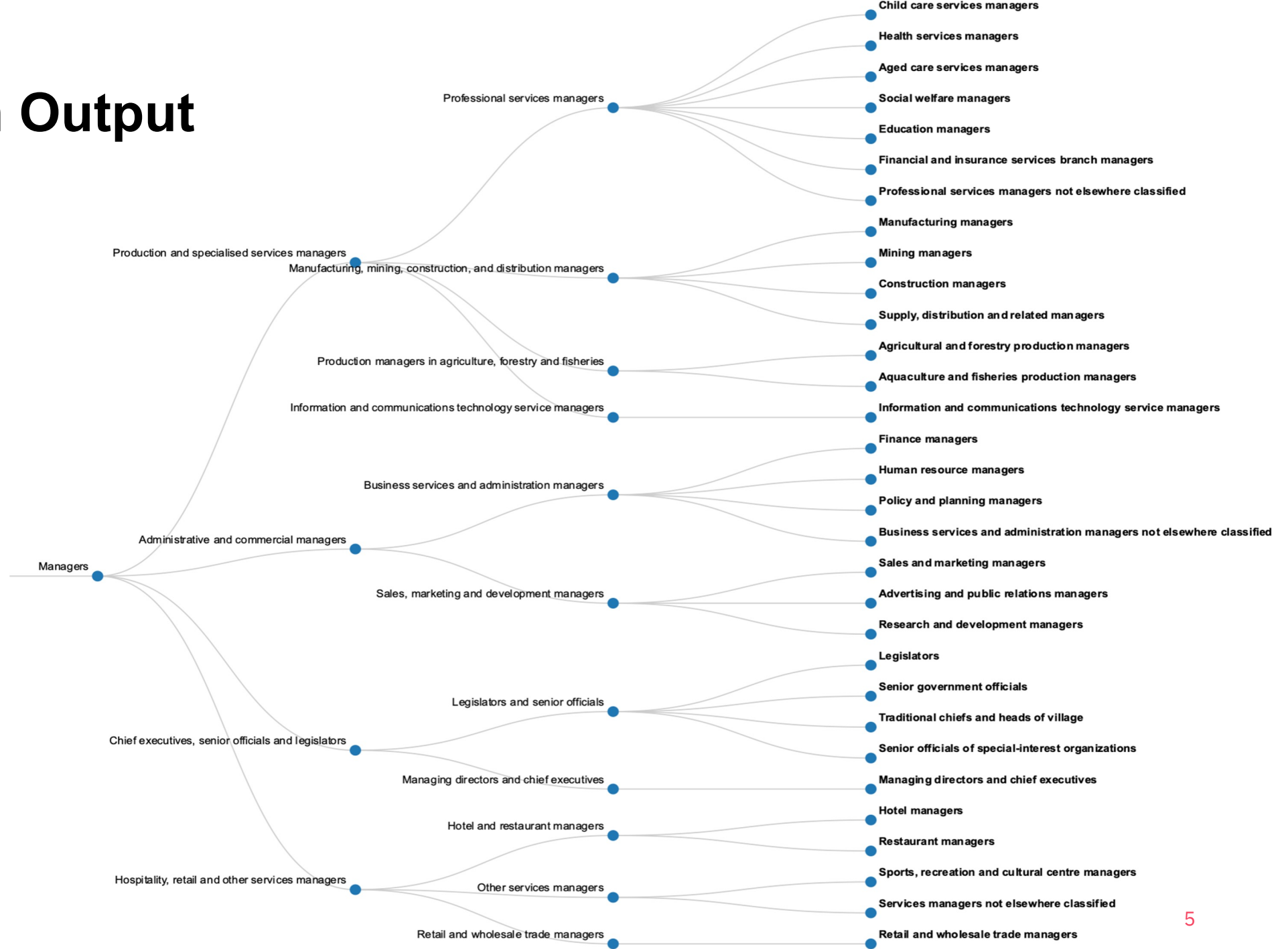 - **Guarantee Explainable information extraction, logging classification methods and relevant features.**

ETF
Working together
Learning for life

European Training Foundation

Data Classification – An example

Job Vacancy

Information extraction

Occupation	Skills
Time	Area
Industry	…

Junior Software Developer

As Junior Software Developer, you will develop excellent software for use in field mapping, data collection, sensor networks, street navigation, and more. You will collaborate with other programmers and developers to autonomously design and implement high-quality web- based applications, restful API's, and third party integration.

We're looking for a passionate, committed developer that is able to solve and articulate complex problems with application design, development and user experiences. The position is based in our offices in Harwell, United Kingdom.

Information extraction

2512 – Software Developer

Skills: develop software, implement web based applications, problem solving, develop user experiences

Harwell, UK

…

Occupation pipeline

Occupation Output

Job Location Recognition

C#, XML and Data Software Developer

Scottish Government ★★★★☆ 40 reviews - Edinburgh
£28,891 - £33,094 a year - Full-time, Part-time, Permanent
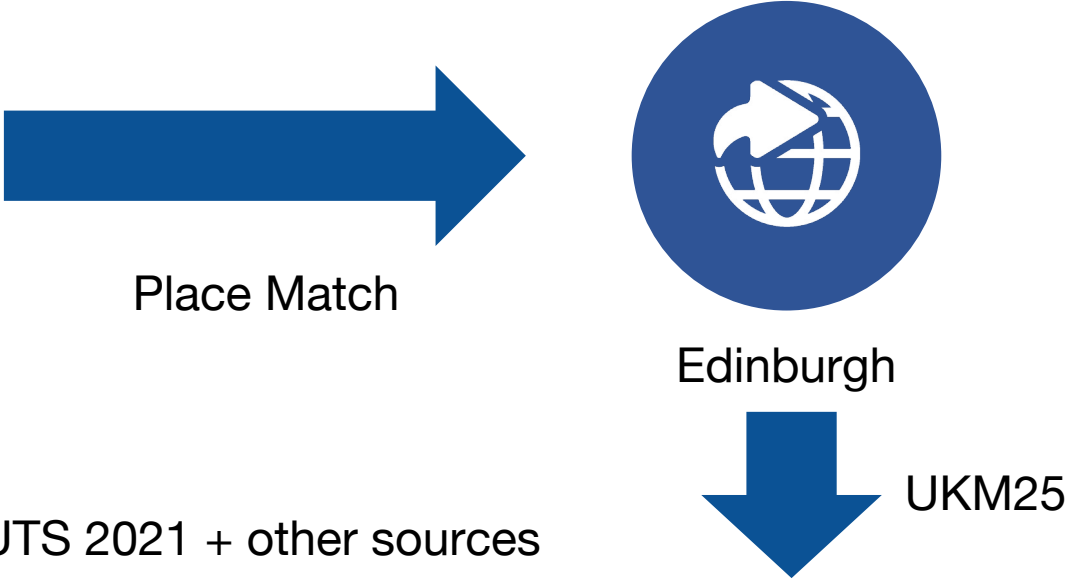
Apply On Company Site

For the full job description and to apply, please visit the Work for Scotland website
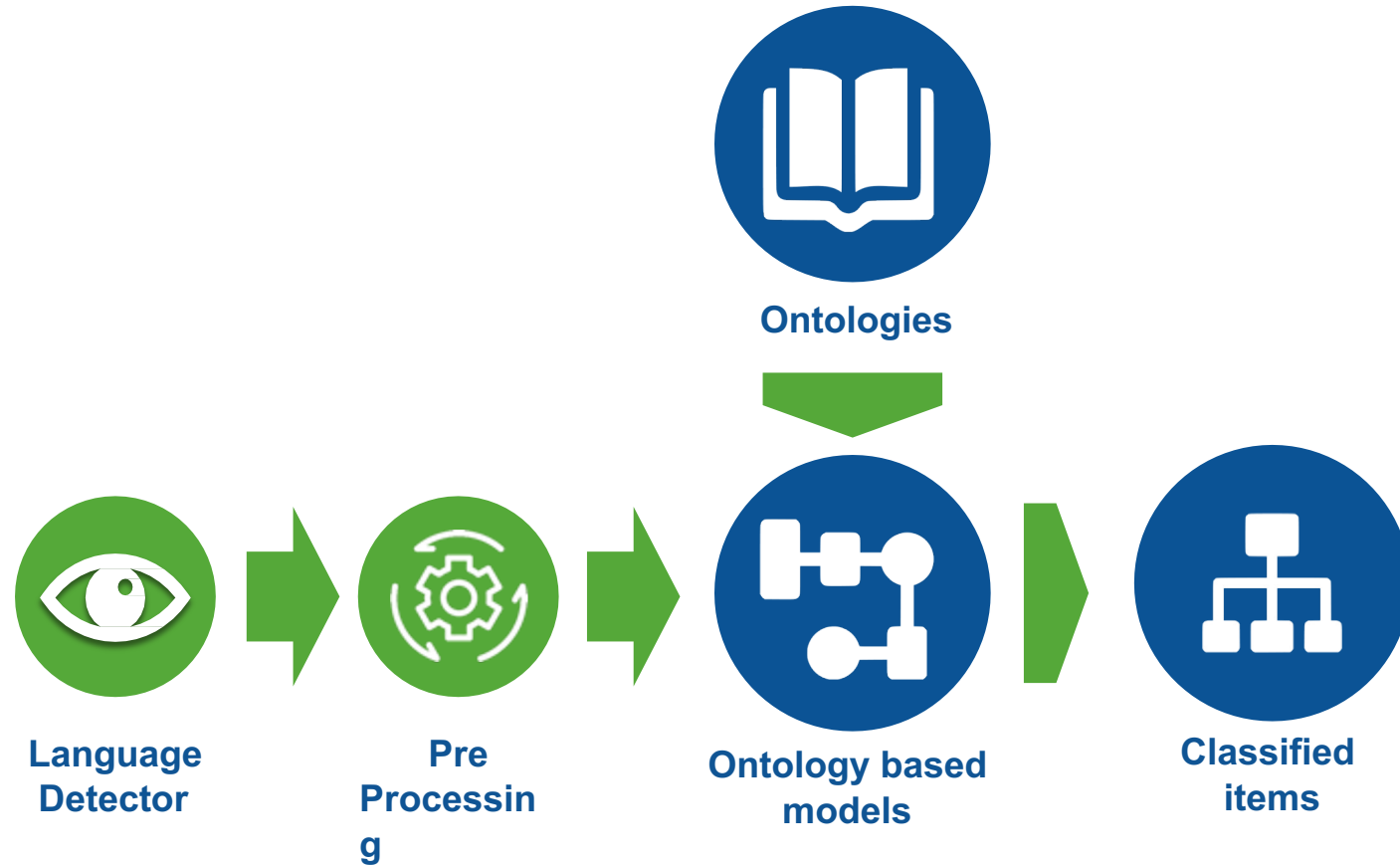Benefits

Pensions

Civil Service pension schemes may be available for successful candidates.

Civil Service   15 days ago - apply on company site
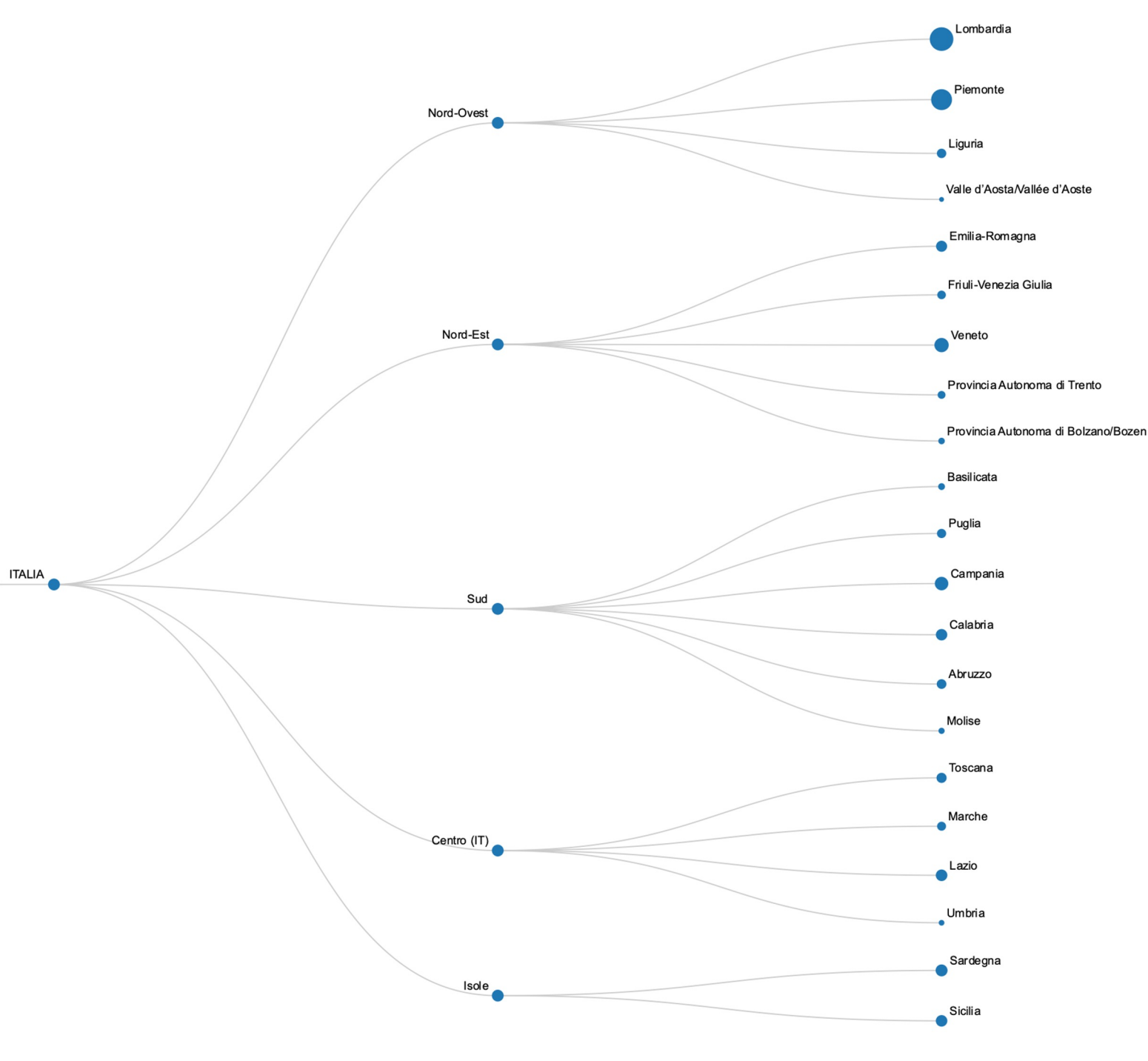
Place Match

Edinburgh

UKM25

NUTS 2021 + other sources

Key	Value
▼ {} (1) {_id : UKM25}	{ 17 fields }
_id	UKM25
lau_1	
lau_2	
idLau_1	
idLau_2	
nutLiv_3	Edinburgh, City of
nutLiv_2	Eastern Scotland
nutLiv_1	SCOTLAND
nutLiv_0	UNITED KINGDOM
idNutLiv_3	UKM25
idNutLiv_2	UKM2
idNutLiv_1	UKM
idNutLiv_0	UK
name	Edinburgh, City of
country	UK
nutsLevel	3.0
idIso31161	UG

6

Job Location pipeline

Ontologies

Language Detector → Pre Processing → Ontology based models → Classified items

Place Output

Skill pipeline

Ontologies

Language Detector → Pre Processing → Ontology based models → Classified items

ETF
Working together
Learning for life
European Training Foundation

(job title, skill, or company)   (city or zip code)   Search   Advanced search

Fabasoft International Services GmbH

Application Developer (w / m)

Linz, Vienna   Solid employment   full-time   Erschienen: vor 12 Tagen

You are the right one, if

- you are interested in designing, implementing and integrating specialized applications and you like to work with high-end development tools and infrastructures

- you like using state-of-the-art web technologies to realize customized software solutions while being able to expertly use standard products / SDKs as well as libraries

- You have a solid computer science education (HTL, FH or University) and solid programming skills with Java and Eclipse

- Ideally, you have already gained experience with SOAP, WSDL and Web Services, as well as object-oriented modeling (UML) skills

We offer you

- working in a dedicated, highly motivated team in a professional environment

- Consistent training and further education as well as broadening horizons by attending international conferences

- interesting development opportunities for Certified Project Manager, Certified Scrum Master etc.

- numerous benefits, benefits and cool events as well

- above-average payment (at least EUR 35,000 gross annual salary, depending on your experience and qualifications also more) and a performance-based bonus system
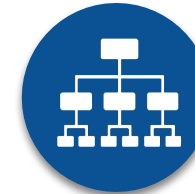
Contact Person: Mag. Dipl.-Ing. Christian Distelberger
E-mail: job@fabasoft.com
Tel: +43 732 606162-0
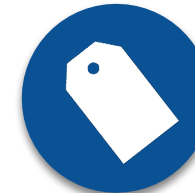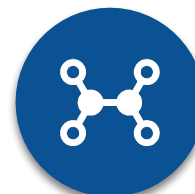
Noise

Disambiguation Text

Alternative labels

Consistent and linkage

ETF
Working together
Learning for life
European Training Foundation

Post a job   For recruiters   Login   Create CV   EN

(job title, skill, or company)   (city or zip code)   Search   Advanced search

Fabasoft International Services GmbH

Application Developer (w / m)

Fabasoft

📍 Linz, Vienna   💼 Solid employment   🕐 full-time   📅 Erschienen: vor 12 Tagen

You are the right one, if

- you are interested in designing, implementing and integrating specialized applications and you like to work with high-end development tools and infrastructures
- you like using state-of-the-art web technologies to realize customized software solutions while being able to expertly use standard products / SDKs as well as libraries
- You have a solid computer science education (HTL, FH or University) and solid programming skills with Java and Eclipse
- Ideally, you have already gained experience with SOAP, WSDL and Web Services, as well as object-oriented modeling (UML) skills

We offer you

- working in a dedicated, highly motivated team in a professional environment
- Consistent training and further education as well as broadening horizons by attending international conferences
- interesting development opportunities for Certified Project Manager, Certified Scrum Master etc.
- numerous benefits, benefits and cool events as well
- above-average payment (at least EUR 35,000 gross annual salary, depending on your experience and qualifications also more) and a performance-based bonus system

Contact Person: Mag. Dipl.-Ing. Christian Distelberger
E-mail: job@fabasoft.com
Tel: +43 732 606162-0

Noise

We offer you

```
<working>
<highly motivated team>
<professional>
<Certified Project
Manager>
<Scrum>
<performance-based>
```

ETF
Working together
Learning for life
European Training Foundation

StepStone

Post a job   For recruiters   Login   Create CV   EN

(job title, skill, or company)   (city or zip code)   Search   Advanced search

Fabasoft International Services GmbH
Application Developer (w / m)

📍 Linz, Vienna   💼 Solid employment   🕐 full-time   📅 Erschienen: vor 12 Tagen

You are the right one, if

- you are interested in designing, implementing and integrating specialized applications and you like to work with high-end development tools and infrastructures

- you like using state-of-the-art web technologies to realize customized software solutions while being able to expertly use standard products / SDKs as well as libraries

- You have a solid computer science education (HTL, FH or University) and solid programming skills with Java and Eclipse

- Ideally, you have already gained experience with SOAP, WSDL and Web Services, as well as object-oriented modeling (UML) skills
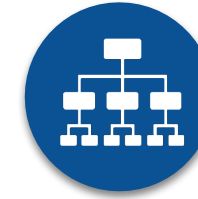
We offer you

- working in a dedicated, highly motivated team in a professional environment

- Consistent training and further education as well as broadening horizons by attending international conferences

- interesting development opportunities for Certified Project Manager, Certified Scrum Master etc.

- numerous benefits, benefits and cool events as well

- above-average payment (at least EUR 35,000 gross annual salary, depending on your experience and qualifications also more) and a performance-based bonus system

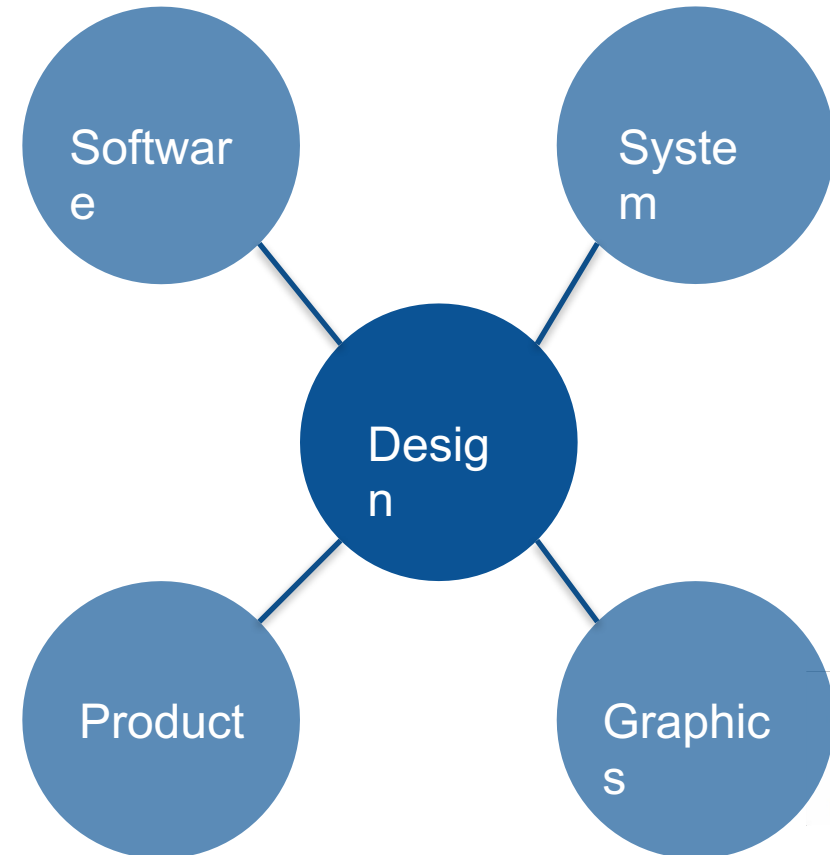**Contact Person:** Mag. Dipl.-Ing. Christian Distelberger
E-mail: job@fabasoft.com
Tel: +43 732 606162-0

Disambiguation Text

`<design> <application>`

Software   System   Design   Product   Graphics

ETF Working together Learning for life — European Training Foundation

12

StepStone

(job title, skill, or company)   (city or zip code)   Search   Advanced search

Fabasoft International Services GmbH

Application Developer (w / m)

📍 Linz, Vienna   💼 Solid employment   🕐 full-time   📅 Erschienen: vor 12 Tagen

You are the right one, if

- you are interested in designing, implementing and integrating specialized applications and you like to work with high-end development tools and infrastructures

- you like using state-of-the-art web technologies to realize customized software solutions while being able to expertly use standard products / SDKs as well as libraries

- You have a solid computer science education (HTL, FH or University) and solid programming skills with Java and Eclipse

- Ideally, you have already gained experience with SOAP, WSDL and Web Services, as well as object-oriented modeling (UML) skills
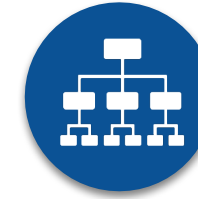
We offer you

- working in a dedicated, highly motivated team in a professional environment

- Consistent training and further education as well as broadening horizons by attending international conferences

- interesting development opportunities for Certified Project Manager, Certified Scrum Master etc.

- numerous benefits, benefits and cool events as well

- above-average payment (at least EUR 35,000 gross annual salary, depending on your experience and qualifications also more) and a performance-based bonus system

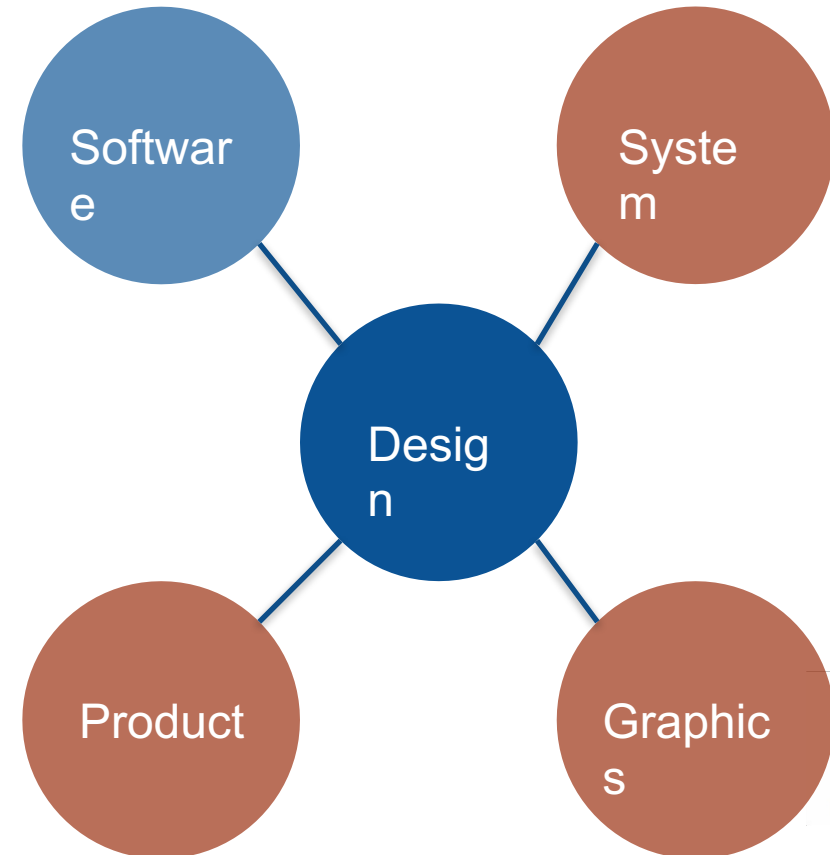**Contact Person:** Mag. Dipl.-Ing. Christian Distelberger
E-mail: job@fabasoft.com
Tel: +43 732 606162-0

Disambiguation Text

`<design> <application>`

Software — System — Design — Product — Graphics

ETF Working together Learning for life
European Training Foundation

13

Post a job   For recruiters   Login   Create CV   EN

(job title, skill, or company)   (city or zip code)   Search   Advanced search

Fabasoft International Services GmbH
Application Developer (w / m)

Linz, Vienna   Solid employment   full-time   Erschienen: vor 12 Tagen

You are the right one, if

- you are interested in designing, implementing and integrating specialized applications and you like to work with high-end development tools and infrastructures

- you like using state-of-the-art web technologies to realize customized software solutions while being able to expertly use standard products / SDKs as well as libraries

- You have a solid computer science education (HTL, FH or University) and solid programming skills with Java and Eclipse

- Ideally, you have already gained experience with SOAP, WSDL and Web Services, as well as object-oriented modeling (UML) skills
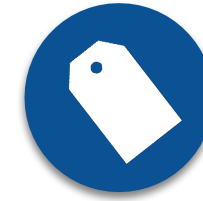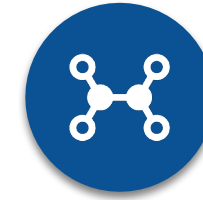
We offer you

- working in a dedicated, highly motivated team in a professional environment

- Consistent training and further education as well as broadening horizons by attending international conferences

- interesting development opportunities for Certified Project Manager, Certified Scrum Master etc.

- numerous benefits, benefits and cool events as well

- above-average payment (at least EUR 35,000 gross annual salary, depending on your experience and qualifications also more) and a performance-based bonus system

Contact Person: Mag. Dipl.-Ing. Christian Distelberger
E-mail: job@fabasoft.com
Tel: +43 732 606162-0

Alternative Labels

`<computer science>`

Computer science

`<ESCO>`
Computer programming

ETF Working together Learning for life
European Training Foundation

🔍 (job title, skill, or company)   📍 (city or zip code)   **Search**   Advanced search

Fabasoft International Services GmbH

Application Developer (w / m)

📍 Linz, Vienna   💼 Solid employment   🕐 full-time   📅 Erschienen: vor 12 Tagen

You are the right one, if

- you are interested in designing, implementing and integrating specialized applications and you like to work with high-end development tools and infrastructures

- you like using state-of-the-art web technologies to realize customized software solutions while being able to expertly use standard products / SDKs as well as libraries

- You have a solid computer science education (HTL, FH or University) and solid programming skills with Java and Eclipse

- Ideally, you have already gained experience with SOAP, WSDL and Web Services, as well as object-oriented modeling (UML) skills

We offer you

- working in a dedicated, highly motivated team in a professional environment

- Consistent training and further education as well as broadening horizons by attending international conferences

- interesting development opportunities for Certified Project Manager, Certified Scrum Master etc.

- numerous benefits, benefits and cool events as well

- above-average payment (at least EUR 35,000 gross annual salary, depending on your experience and qualifications also more) and a performance-based bonus system

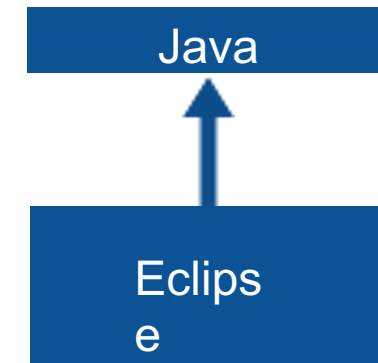**Contact Person:** Mag. Dipl.-Ing. Christian Distelberger
E-mail: job@fabasoft.com
Tel: +43 732 606162-0

Consistent and linkage

`<java> <eclipse>`

Java

↑

Eclipse

Dataset Model

2 Tables

FT Document

1 Row for each:
- General_ID - Key
- OJV
- Source
- Place

Why? Because, for each OJV, we can detect a multi-place Vacancy
  e.g. «Software Developer in London / Liverpool»

FT Skill Analysis

1 Row for each:
- General_ID - Key
- OJV
- Source
- Place
- Skill

Why? Because, for each OJV, we can obviously detect multiple skills
 e.g. «Software Developer in London / Liverpool, with customer orientation culture, that speaks english and tolerates stress»

16

ETF
Working together
Learning for life

European Training Foundation

Dataset Model

So, to have the number of job vacancies you always have to compute an unique count by General_ID

ETF
Working together
Learning for life

European Training Foundation
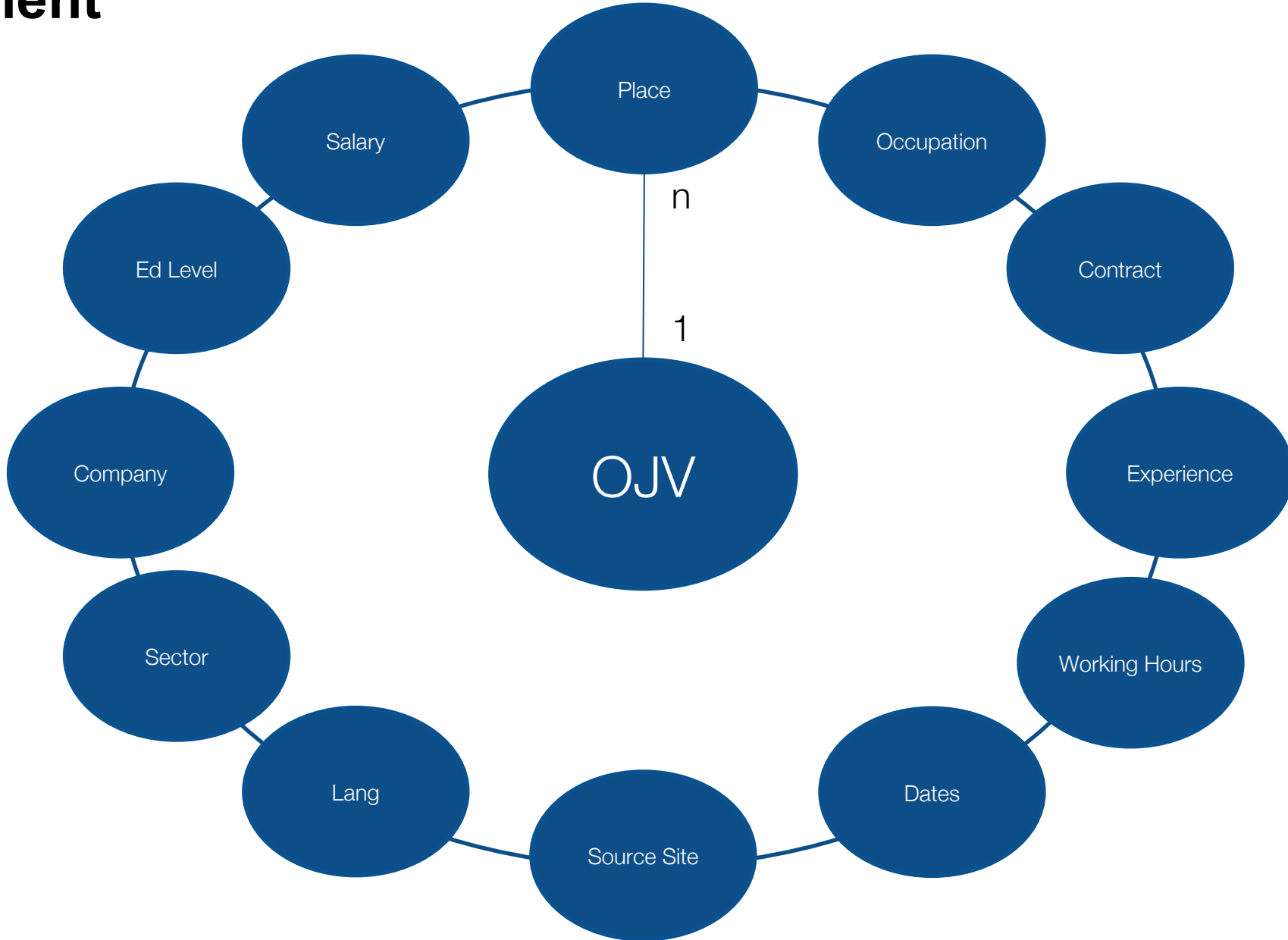
FT Document

FT Skill Analysis
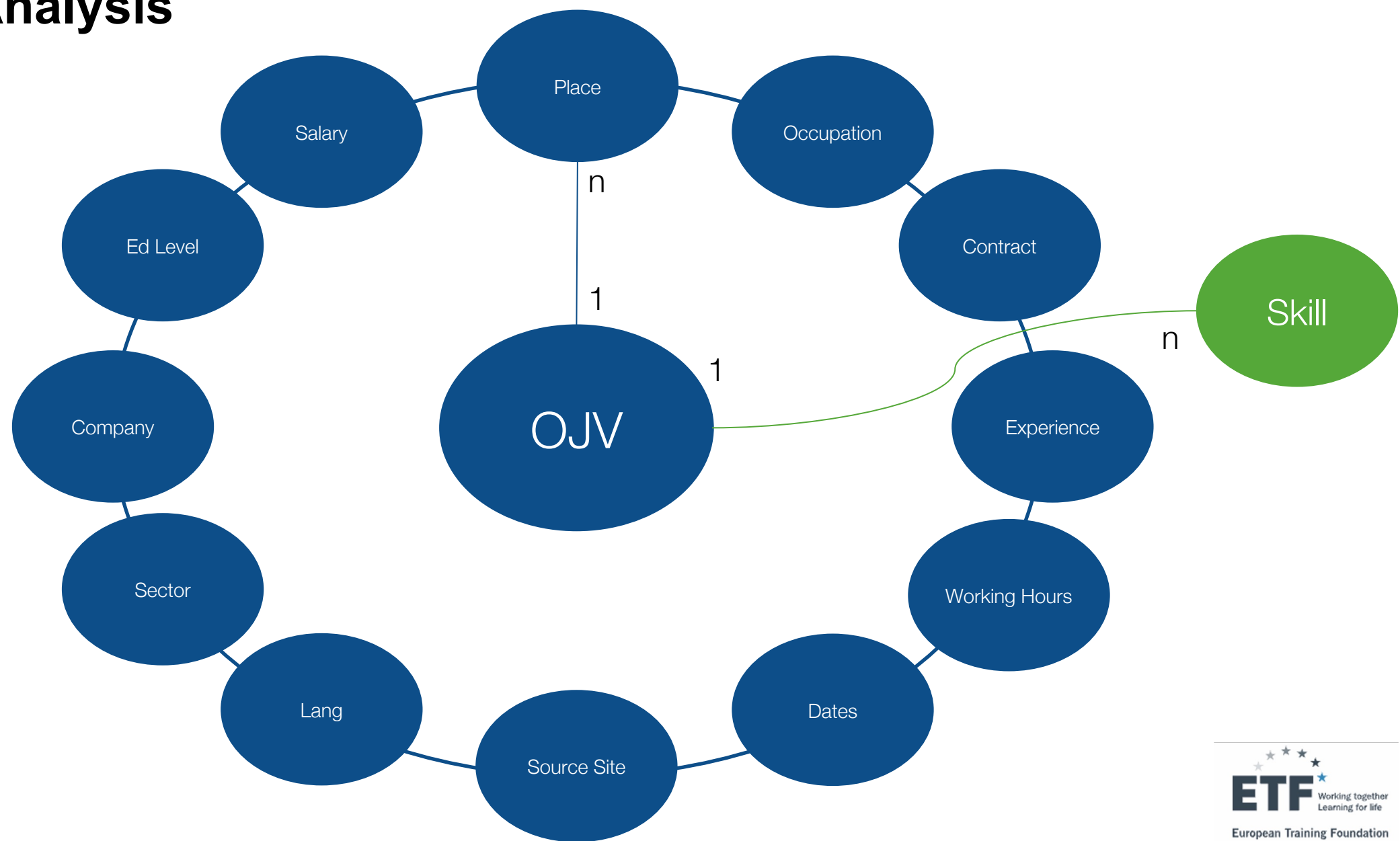
Tools

AWS Athena Console

Live session with python and SQL