



# European Data Science Academy

---

[edsa-project.eu](http://edsa-project.eu)  
[@edsa\\_project](https://twitter.com/edsa_project)

# What is EDSA about

**ESTABLISH A EUROPEAN  
COMMUNITY OF EDUCATORS,  
LEARNERS, PRACTITIONERS, AND  
POLICY MAKERS ON DATA SCIENCE**



# Who we are



UNIVERSITY OF  
Southampton

TU/e Technische Universiteit  
Eindhoven  
University of Technology



ODI open  
data  
institute



ideXlab

Fraunhofer  
IAIS

PERSONTYLE  
DATA SCIENCE CENTRE OF EXCELLENCE

# What does EDSA do

## Landscaping

- Surveys
- Interviews
- Dashboards

## Curriculum and courseware

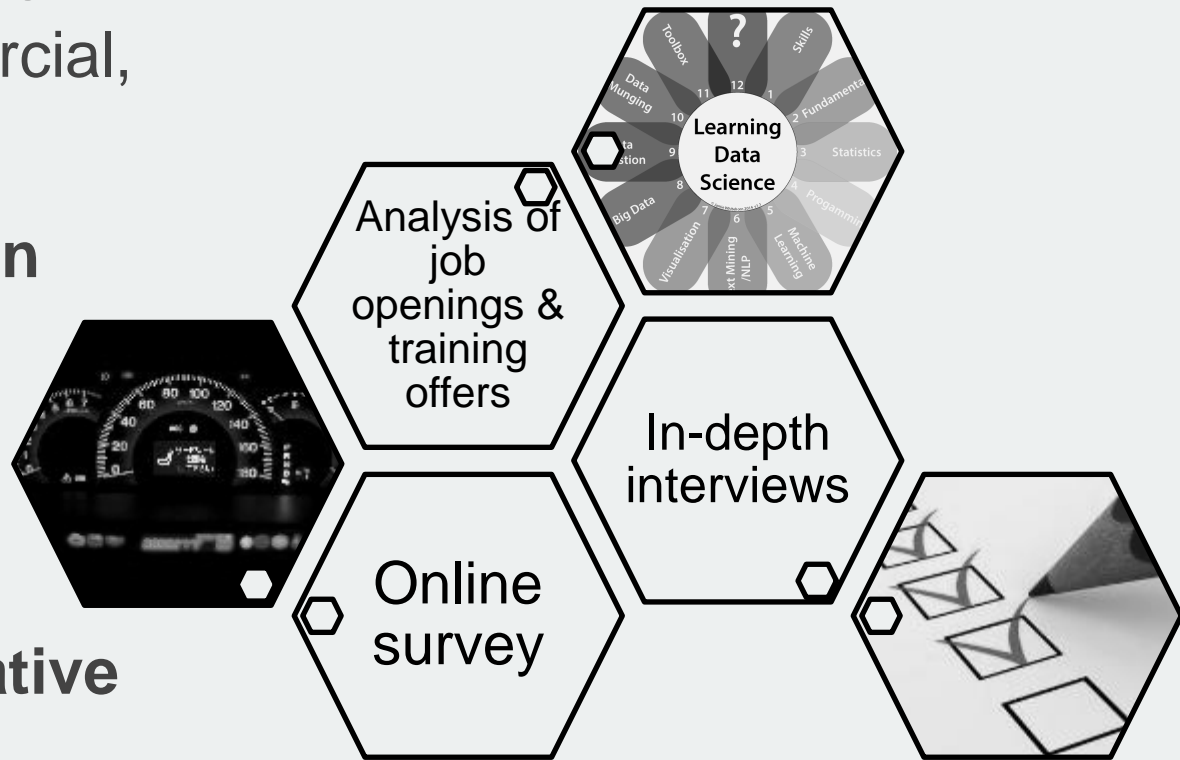
- Modular, media-rich, multiple languages
- Core, domain-specific, and technology-specific topics

## Courses and learning analytics

- Video lectures
- Professional training
- MOOCs
- eBooks

# Our landscaping approach

- **Aim: study data science demand**
  - In EU28 countries
  - Sectors: commercial, government, academia
- **Demand: defined in terms of skills**
  - Essential and desirable
  - Topics and technology
- **Methods: Quantitative and qualitative, integrated within a common framework**



# Online survey

<http://edsa-project.eu/resources/survey/>

For data scientists and their managers

For EU organizations of different sizes and sectors

How important is each of the following skills for a data scientist?

Drag the elements to rate their importance in your sector (either **essential**, **desirable** or **not required**). You can also **add** to each category.

Essential	Desirable	Not required
<b>Add</b>	<b>Add</b>	<b>Add</b>
Big data	Open source tools and concepts	Data collection and analysis
Advanced computing and programming	Data interpretation and visualisation	Math and statistics
Machine learning and prediction	Business intelligence and domain expertise	

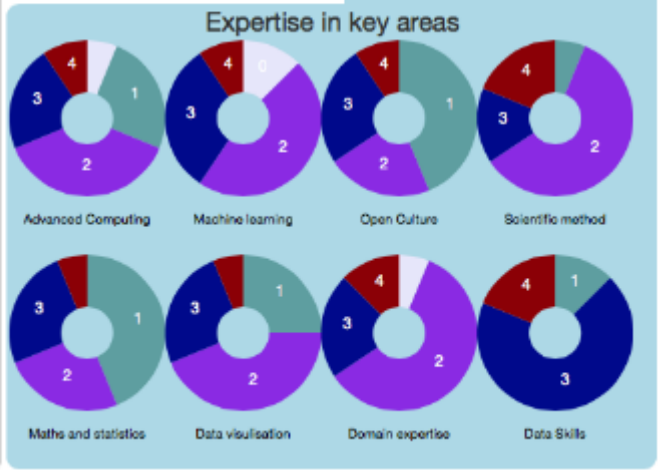
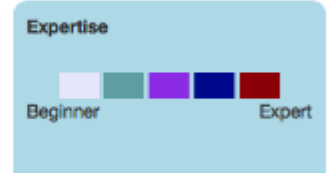
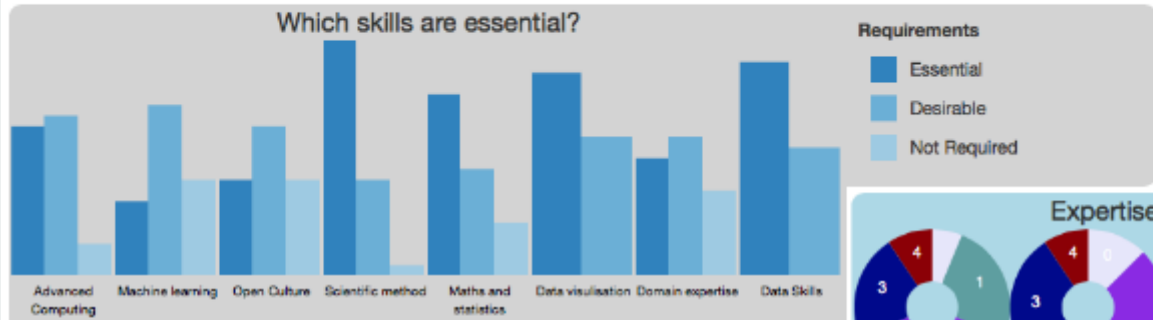
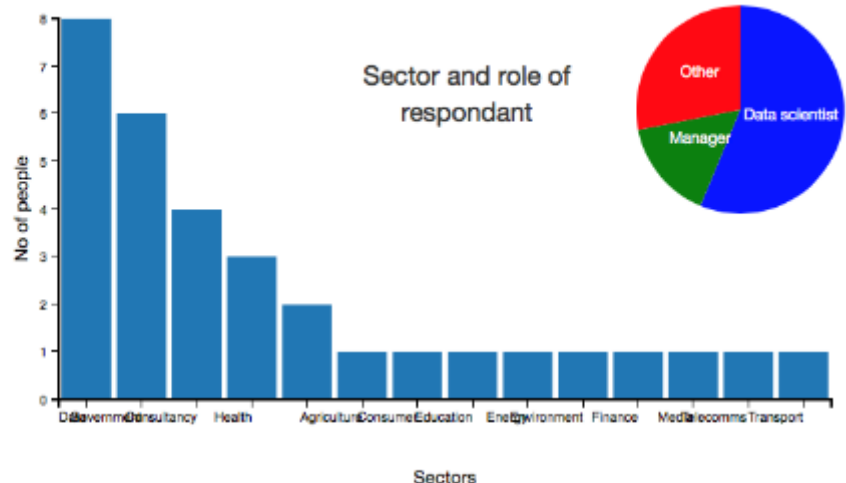
What training delivery methods do you consider to be important factors in successful training?

Essential	Desirable	Not required
<b>Add</b>	<b>Add</b>	<b>Add</b>
Face to face training	Webinars	eLearning
Translated from English	Tailored to sector	Accredited
Uses non-open, non-free software	Coaching	Assessed
	Internal assignments	

Which data technologies, tools and languages would you like to see included in data science training?

[AWS](#) [Spark](#) [Hadoop / MapReduce](#) [MongoDB](#) [Open Refine](#) [QMiner](#) [Apache Flink](#) [Apache Storm](#) [ProM or Disco](#) [NoSQL / Cassandra](#) [R](#) [Python](#) [Javascript / JQuery](#) [D3 / nvD3](#) [Java](#) [z-scores](#)

# Survey response data



# Analysis of job openings

<http://edsa-project.eu/resources/dashboard/>

- Visual analysis of job postings
  - Multiple languages, sectors, types of jobs, data sources
- We assess
  - Skills required in each country
  - Number of jobs per country and per area (at city level)
  - Skills clusters
- *We match these against training courses*



# Processing pipeline

## Data collection

GitHub

StackOverflow

*LinkedIn*

*Monster*

*Kaggle*

## Pre-processing

Map to schema

Semantic annotations

Duplicate elimination

## Analysis

Extract skills

Extract job type

Extract organization and sector

Cross-lingual post clustering

## Storage and maintenance

Dashboard data store

SPARQL endpoint

Visual interface

# Skills and locations

**Filter by Location** | Current filter: Munich

- London
- Dublin
- Berlin
- Brussels\*
- Paris
- Amsterdam
- Munich**
- Warsaw
- Madrid
- Prague
- Barcelona
- Luxembourg\*
- Bucharest
- Budapest

**Filter by Skill**

- python
- java
- sql
- big data
- r
- performance
- computer-science
- data-mining
- e-commerce
- nosql
- project-management
- data-management
- web
- administrator
- databases

Select a Country **Job count: 57**

Range of Jobs

- 0-24
- 25-47**
- 48-66
- 67-142
- 143-184
- 185-340

**Detail for Selection (job count: 157)**

Italy	Affli*	.net	1	VITROCISET*
Italy	Milan*	r	1	Atmen S.r.l.*
Italy	Turin*	database	1	RGI S.p.a.*
Italy	Trento	amazon-web-services	1	SkipassCo srl
Italy	Genoa	analytics	1	Ikrix*
Italy	Zola Predosa*	analytics	1	Yoox*
Italy	Vicenza*	architecture	1	Atmen S.r.l.*
Italy	Livorno*	assessment	1	Atmen S.r.l.*

**Filter by Location**

- Milan\*
- Rome\***
- Bologna\*
- Genoa
- Florence\*
- Turin\*
- Novara\*
- Treviso\*
- Vicenza\*
- Bari\*
- Pisa\*
- Trento
- Ancona\*
- Are\*
- Bergamo\*

**Filter by Skill**

- engineering
- image
- computer-science
- configuration
- finance
- product
- software
- engineer
- microsoft
- modeling
- object
- project-management
- sales
- software-development
- dynamic



# EDSA Ambassadors Program

Help train the next generation of data scientists

Be part of the 1<sup>st</sup> EU-wide initiative on data science

Ambassador

Certified courses

Trainer

Trainee





# European Data Science Academy

---

[edsa-project.eu](http://edsa-project.eu)  
[@edsa\\_project](https://twitter.com/edsa_project)