

EUROPEAN COMMISSION EUROSTAT

First hand demo of winning project 1<sup>st</sup> EU Big Data Hackathon

# **EU Jobs and Skills Explorer**

#### Belgrade, Serbia, 12<sup>th</sup> July 2017.



## Agenda

- 1. Team
- 2. About IN2data (what we do)
- 3. EU Big Data Hackathon Executive Summary
- 4. EU Big Data Hackathon Timeline
  - Team Selection Pre-Hackathon Activities
  - Competition
  - Policy Question
  - Platform Selection
  - Solution/Output
  - Panel Members
- 5. Prototype Live Demo
- 6. Prototype Credentials
- 7. Q&A

### TEAM

# About Us



Igor Kaluđer, PhDc igor.kaludjer@in2data.eu

Leo Mršić, PhD leo.mrsic@in2data.eu

Robert Kopal, PhD robert.kopal@in2data.eu

Milica Mićić, PhD milica.micic@in2.rs

Average 18+ years of work experience per expert, active in business/EDU/GOV community, PhD/PhDc level in advanced analytics, Professors/lecturers, 50+ scientific papers, 15+ books, 100+ conferences/lectures, data economy professionals, data monetization experts...

#### IN2DATA – WHAT WE DO



## About IN2data Company

#### Unique team, 8 team members (data scientists), Data economy oriented, Vendor agnostic...







# Data Products / 60+



# **Business Model**





#### **Education and Training**

Constant model(s) improvement and knowledge transfer

#### **Unlock business value**

By having analytics platform in place, base has been set for further model development

#### Data Products + Visualization

Industry/cross-industry data products

#### **Analytics Platform**

All relevant data integration on platfrorm ready to build analytical models

### EU BIG DATA HACKATHON EXEC SUMMARY

## EU Big Data Hackathon



Google Big Query Tableau Cloud Live demo

### FIRST EU BIG DATA HACKATHON IN DETAIL

### 1st EU Big Data Hackathon



DRŽAVNI ZAVOD ZA STATISTIKU REPUBLIKE HRVATSKE CROATIAN BUREAU OF STATISTICS

#### OBJAVA REZULTATA JAVNOG POZIVA ZA SUDJELOVANJE NA

EUROPSKOM BIG DATA HACKATHONU

Izuzetno nam je zadovoljstvo obznaniti da je javni poziv Državnog zavoda za statistiku za prijavu natjecateljskih timova za sudjelovanje na Europskom Big Data Hackathonu izazvao veliki interes stručne javnosti te smo zaprimili mnoštvo nadasve kvalitetnih pojedinačnih i timskih prijava.

Nakon pomnog pregleda svih pristiglih prijava, stručni ocjenjivački sud kojeg su činili Saša Cvetojević, direktor Insaco d.o.o., Luka Ljubičić, pomoćnik ravnatelja za informatiku Hrvatskog zavoda za mirovinsko osiguranje te Marko Krištof, ravnatelj Državnog zavoda za statistiku, donio je odluku da će na Europskom Big Data Hackathonu u ožujku 2017., Hrvatsku predstavljati tim kojeg čine doc. dr. sc. Robert Kopal, dr. sc. Leo Mršić i mr. sc. Igor Kaluđer.

Zahvaljujemo svim pojedincima koji su iskazali svoj interes i motiviranost za sudjelovanje na Europskom Big Data Hackathonu te moramo naglasiti da izbor nije bio nimalo lak. Pred odabranim timom nalazi se veliki izazov te im želimo puno sreće!



Zagreb, 11. studenoga 2016.

#### Team Croatia



#### Core team Igor Kaluđer, MSc Leo Mršić, PhD Robert Kopal, PhD + IN2data team (Milica Mićić, PhD)



**NSI** Maja Pekeč Vesna Lipavić Mario Gavrić



Croatian Employment Service (EURES) Igor Vojnić Vilma Mostahinić



Faculty of Law Teo Matković





## Competition

First Big Data hackathon organized by European Commission and EUROSTAT

- ✓ 22 teams (65 experts)
- ✓ 21 countries
- ✓ 2 days & 3 nights
- ✓ 6 minutes pitch per country



#### Data Sets

4.7 mil EURES CVs 35.0 mil EURES job ads

7.4 mil web scraped job ads

**1.0 mil** records from EU Statistics on Income and Living Conditions Survey (SILC)

**0.2 mil** records from OECD Programme for the International Assessment of Adult Competences Survey (PIAAC)

1.4 mil records from EU Labour Force Survey (LFS)

...other surveys

### **Policy Question**



### **Policy Question**

How would you support the design of policies for reducing mismatch between jobs and skills at regional level in the EU through the use of data?

#### Demonstrate your work through a tool that performs data analytics on the provided data and potentially additional data sources.

Your proposal should take into account various approaches in matching "demand for" and "supply of" skills either by providing relevant training, outsourcing or hiring employees. The proposal should be supported by relevant data, statistical analysis and visualisation. You are also invited to establish links between various data sets and to identify and use other data sources beyond those currently at hand. The tool should integrate or at least consider the aspect of data quality for the analysis and visualisation. The tool could interactively provide answers to specific questions as shown in the following <u>examples</u>:

- □ How can mismatches between skills demand and skills supply in the labour market be measured? Which regions have reported the biggest mismatches?
- □ How can mismatches between skills demand and skills provided by education system be measured?
- U Which regions provide higher training opportunities (formal-informal education, on the job training)? How can this be measured?
- Can you identify regions with high/low labour mobility (senders/receivers of skills)?
- □ Which regions have the highest potential to reduce skills mismatch through intra-EU labour mobility?
- How business environment (legal, financial, economic, labour, entrepreneurship culture, etc.) is linked to skills demand/supply/ (mis)match in the EU regions?
- What are the most (and least wanted) jobs at regional level in the EU? Can the top (bottom) 10 jobs be identified at regional level in the EU?
- □ What are the most (and least wanted) skills at regional level in the EU? Can the top (bottom) 10 skills be identified at regional level in the EU?

## **Skill Decomposition & Matching**



| PMP CERTIFIED PROJECT AND  | ORGANIZATIONAL MANAGER   | 2.84  |  |  |  |
|--|--|---|--|--|--|
| NewYork, NY 555-555-5555 Jaravogiernal.com   |  | 11  |  |  |  |
| Key Skills Assessment  | Certi  | ifications &  |  |  |  |
|  | Men  | nberships   |  |  |  |
| <ul> <li>Strategic Management - Experienced in vendor mi<br/>strategic financial analysis to increase efficiencies a</li> </ul>      | in agement, contract negotiations and<br>nd reduce costs while maintaining solid   | Comment Deductored  |  |  |  |
| rapport with business partners.  |  |   |  |  |  |
| <ul> <li>Team Leadertrap and Budget Haragement-<br/>cross functional teams and collaborating with s</li> </ul>                       | John Bravo 🦳 👘   |   |  |  |  |
| and starting up new business locations while or  | DND CERTIFIED DDO JECT AND ODCANIZATIONAL MANAGER  |   |  |  |  |
| <ul> <li>Process Improvements – Devised and implem<br/>internal controls to strengthen operations, ince<br/>satisfaction.</li> </ul> | MP CERTIFIED PROJECT AND C<br>New York, NY 555-5555 (provogenal.com  |   |  |  |  |
| Professional Experience  |  |   |  |  |  |
|  | Key Skills Assessment  | Certifications &  |  |  |  |
| Manager and Owner  |  | Memberships   |  |  |  |
| AMERICAN FAMILY INSURANCE  | <ul> <li>Strategic Management - Depertenced in vendor mana<br/>invester Respond analysis to investor and</li> </ul>  | gement, contract negotiations and<br>advancementable examining solid  |  |  |  |
| Established a successful insurance business from d   | rapport with business partners.  | Project Management Professional     IRMET Certification   |  |  |  |
| growing the business and implementing policies at<br>Constant violations to train and manage bases me                                | <ul> <li>Team Leadership and Budget Management- Experim<br/>core functional teams and callaboration with remi</li> </ul>   | a is backet and is differentiation  |  |  |  |
| <ul> <li>Received the highest esteemed award, Washing</li> </ul>   | and starting up new business locations while overs   | John Bravo  | -  |  |  |
| to American Family agents for outstanding achi<br>business and tala conduction   | <ul> <li>Process Improvements - Devised and implement<br/>internal controls to strengthen operations, increase</li> </ul>  | Join Diavo  | 2.00   |  |  |
| <ul> <li>Developed guidelines for personnel evaluations</li> </ul>   | utidaction.  | PMP CERTIFIED PROJECT AND ORGANIZATIONAL M  | ANAGER 😒   |  |  |
| procedures.  |  | New York, NY 555-555-5555 (bravogivenal.com   |  |  |  |
| <ul> <li>Maintained encient team structures and evaluation</li> </ul>  | Professional Experience  |   |  |  |  |
| Office Manager/Financial Coordinator/M   |  | 1   |  |  |  |
| WESLEY JOHNSON, DDS  | Manager and Owner  | Key Skills Assessment   | Certifications &   |  |  |
| Managed the day-to-day operations filling many ro  | For the second s | Strategic Hanagement - Doerierced in under management, contract registrations and   | memoerships  |  |  |
| bottom line. Developed and implemented compret   | growing the business and implementing policies and p   | strategic financial analysis to increase efficiencies and reduce costs while maintaining solid  | Project Management Professional  |  |  |
| resources (HR) responsibilities.   | <ul> <li>Received the highest esteemed award, Washington</li> </ul>  | Team Leadership and Budget Hanagement- Dipartise in leading and building cohesive   | (PMP) Certification<br>Project Management Institute (PMI)                                  |  |  |
| <ul> <li>Utilized strategic planning methodologies and</li> </ul>  | to American Family agents for cutstanding achieve<br>business and sala randoction  | cross tunctionaritiants and consolvations while overseeing staff and managing budgets.  | Member<br>Seattle Dumber of Commerce   |  |  |
| <ul> <li>Measured programs and processes for strength</li> </ul>   | <ul> <li>Developed guidelines for personnel evaluations, st</li> </ul>   | <ul> <li>Process improvements – Deviced and implemented processes, procedures, systems and<br/>internal controls to strengthen operations, increase productivity and enhance customer</li> </ul>    | Member   |  |  |
| changes as needed.<br>• Errors dated college and strategic place and im  | <ul> <li>Maintained efficient team structures and evaluated</li> </ul>   | unidation   | Association Mamber   |  |  |
| <ul> <li>remnance powers and strategic plans and in<br/>initiatives.</li> </ul>  |  | Professional Experience   | <ul> <li>Women in Susiness Wember</li> </ul>   |  |  |
|  | Office Manager/Financial Coordinator/Mark  | FIGURATION Experience   | Awards & Recognitions  |  |  |
| Licensed Realtor   | WESLEY JOHNSON, DOS  | Manager and Owner 2010 - 2013   |  |  |  |
| SALT LAKE BOARD OF REALTORS  | Managed the day to day operations filling many roles t<br>bottom line. Developed and implemented comprehen   | AMERICAN FAMILY INSURANCE   | <ul> <li>Washington State Agent of the<br/>Month. June 2013</li> </ul>                     |  |  |
| <ul> <li>Achieved \$3.5M in sales, Rockie of the Year, and</li> <li>Consecutively exceeded weeks and re-</li> </ul>                  | providing productive administrative management invo<br>tencement (N2) representibilities   | Established a successful insurancebusiness from the ground up, strategically and innovatively<br>growing the business and implementing policies and programs to increase company profits.           | <ul> <li>Received the highest entermed</li> </ul>  |  |  |
| · ····································   | Utilized strategic planning methodologies and gos  | <ul> <li>Consted systems to train and manage team members in achieving agency objectives.</li> <li>Received the highest externed agent. Maximutan State Agent of the Nardy in 2013 gives</li> </ul> | agents for outstanding   |  |  |
|  | <ul> <li>Increasing productivity and profitability.</li> <li>Measured programs and processes for strengths an</li> </ul>   | ts American Family agants for outstanding achievement in sussignition for an overall mix of   | achievement invecogrition for a<br>overall mix of business and sale                        |  |  |
|  | charges as needed.<br>Consolitated tolking and strategic above of instance   | Developed guidelines for personnel evaluations, staff advancement and recruitment.  | production.  |  |  |
|  | <ul> <li>remnance packs and stategic plans and impleministatives.</li> </ul>   | procedures.<br>Maintained efficient team structures and evaluated performance based on analytics.   | All American Club Read   |  |  |
|  |  | Office Manager Property Construction Manhattan Provider 2003  | <ul> <li>Angely extended, outstanding<br/>achievement recognition in the</li> </ul>        |  |  |
|  | Licensed Hearton   | WEST FY JOHNSON, DOS  | Next year of sales in 2003, 2012<br>and ahead of schedule to receive                       |  |  |
|  | <ul> <li>Achieved \$3.5M in tables, Rookie of the Year, and To</li> </ul>  | Managed the day to day operations filling many roles to support growth and add to the   | Ror 2013.  |  |  |
|  | <ul> <li>Consecutively exceeded yearly and monthly compared</li> </ul>   | bottom line. Developed and implemented comprehensive sales and marketing strategies while<br>providing productive administrative management involving business, financial and human                 | All Star Top Agents Club   |  |  |
|  |  | <ul> <li>Information (vie) responsibilities.</li> <li>United students signature methodologies and and other to devit sometimes while</li> </ul>   | <ul> <li>Awarded to a limited number of<br/>top producers in the state.</li> </ul>         |  |  |
|  |  | increasing productivity and profitability.  |  |  |  |
|  |  | <ul> <li>resource programs are provident or complete and wearbitted and implemented<br/>changes as readed.</li> </ul>   | <ul> <li>Life Insurance Honors Club and Life<br/>Diamond-Award</li> </ul>                  |  |  |
|  |  | <ul> <li>Formulated poticies and strategic plans and implemented procedures and growth<br/>initiatives.</li> </ul>  | <ul> <li>Received while meeting American<br/>Family's expectations in sales for</li> </ul> |  |  |
|  |  |   | Home, Buto and Business  |  |  |
|  |  | Licensed Realtor 1999 - 2007  | expectation in Life insurance sale   |  |  |
|  |  | <ul> <li>Achieved SISM in sales, Rookie of the Year, and Top New Realtyr and Female Agent.</li> </ul>   |  |  |  |
|  |  | <ul> <li>Conservatively exceeded yearly and monthly company sales peak.</li> </ul>  | Languages  |  |  |
|  |  |   |  |  |  |

## Output

- ✓ Platform selection
- ✓ Data preparation
- ✓ Data Storage
- ✓ Data Optimization
- ✓ Project Scope Definition
- ✓ Data Processing / Analytics
- ✓ Visualization
- ✓ Presentation Bundle

#### Google BigQuery

| COMPOSE QUERY                  |   | Table De      | tails: ( | OCCUPATIO           | N_DEM      | AND      |
|--------------------------------|---|---------------|----------|---------------------|------------|----------|
| uery History                   |   | Schema        | Details  | Preview             |            |          |
| ob History                     |   | Description   | •        |                     |            |          |
| doc_temp                       | ^ | Description   | •        |                     |            |          |
| ESJS                           |   | Describe this | s table  |                     |            |          |
| ft_document_en                 |   | Table Info    |          |                     |            |          |
| ft_document_en_3               |   | Table ID      | e        | u-hackathon:cedefo  | DD OCCUPAT | ON DEMAN |
| ft_skill_profession_en         |   |               | -        |                     |            |          |
| ft_skill_profession_experience |   | Table Size    | 2        | 45 GB               |            |          |
| ft_skill_profession_new_en     |   | Number of R   | Rows 9   | 32,809,512          |            |          |
| ft_skill_profession_requiremen |   | Creation Tim  | ne M     | ar 13, 2017, 5:22:5 | 5 PM       |          |
| joined_skills_doc_tmp          |   | Leat Nedifie  |          | 40, 0047, 5:00:5    | 5 DU       |          |
| occ_dem_tmp                    |   | Last wodifie  | iu M     | ar 13, 2017, 5.22.5 | 5 FM       |          |
| OCCUPATION_DEMAND              |   | Data Locatio  | on U     | S                   |            |          |
| OCCUPATION_VOLATILITY          |   | Labela        | N        | one                 |            |          |
| out1                           |   | Lapels        |          |                     |            | Edit     |
| SKILL DEMAND BY MONTH          |   |               |          |                     |            |          |

### Platform



### PROTOTYPE

### Solution

Live cloud/web based application, visual, interactive dashboards (region maps included)

Key areas (4):

- 1) Skill supply Skill Supply Skill Clustering Obsolete Skills
- 2) Skill Demand Skill Demand Demand Prediction Occupation Complexity
- 3) Skill Gap Skill Gap Occupation Migration Socio-economic Indicators
- 4) Policy Actions Qualifications Skill & Qualification Gap Workforce Migration



## Skill supply



#### **Skill Clustering**



#### **Occupation Migration**

#### Ξ

#### **Occupation Migration**

#### **Desired** Occupation Clerical Craft and Elementa Plant and Service Skilled ag Technicia Armed forces oc.. support .. related t.. ry occup.. Managers machine .. Professi, and sale.. ricultura.. ns and a.. **Current Occupation** Armed forces occupations 150 362 363 198 858 420 889 974 982 Clerical support workers 34 66,309 4,916 11,034 12,105 5,785 53,706 25,241 2,048 35,745 Craft and related trades ... 7,648 4,864 8,576 1,726 10,680 11,248 2,232 Elementary occupations 3 8.290 6.433 18,990 1.763 4.654 10.666 12.859 2.431 7,484 Managers 66 12,186 3,590 2,723 72,003 2,819 69,472 11,232 1,354 35,418 Plant and machine operat.. 7,278 1,234 25,942 5,524 5,417 1,288 4,506 6,028 6,189 Professionals 41,741 12,856 69,488 6,696 4,815 140,465 84 10,170 33,845 55,914 Service and sales workers 99 36,321 10,407 19,638 11,940 9,106 74,195 2,864 38,501 Skilled agricultural, fores.. 1,014 1,779 1,658 2,378 984 3,918 1,415 730 Technicians and associate.. 155 41,735 15,748 12,143 36,680 9,289 149,531 32,248 2,754 161,189

#### Occupation Level

● 0○ 1

O 2 O 3

Reporting Month

November 2016

#### Nuts0 Name

✓ (AII) ~ ▼ BELGIQUE-BELGIË ✓ ČESKÁ REPUBLIKA ✓ DANMARK ✓ DEUTSCHLAND Y Z EECTI Nuts1 Name ✓ (AII) ~ ALAND ✓ ALFÖLD ÉS ÉSZAK ✓ ATTIKH (ATTIKI) ✓ BADEN-WÜRTTE.. 

Nuts2 Name

✓ (AII)
 ✓ Abruzzo
 ✓ Åland

~

✓ Alentejo

#### Socio-economic Indicators



#### **Occupation Complexity**

#### Ξ

#### **Occupation Complexity**

| Occupations                              | Skills                    | Occupation Leve |
|--|---------------------------|-----------------|
| Physical and engineering science tec 3   | ×                         | 01              |
| Engineering professionals (excluding     | 3D modelling              | ∩ O 2           |
| Sales and purchasing agents and bro 3    | Accounting (adult educati |                 |
| Client information workers 3             | Accounting / auditing     | 0 4             |
| Administrative and specialised secre 3   | Accounting responsibility | Skill Level     |
| Shop salespersons 3                      | Active Server Pages (ASP) | 00              |
| Material-recording and transport cle 3   | Actuarial science         | 01              |
| Sales, marketing and public relations 3  | Administration of wages   | ② 2             |
| Manufacturing, mining, construction, 3   | Adobe Photoshop           | 03              |
| Administration professionals 3           | Advanced Business Applic  |                 |
| Software and applications developer 3    | Advertising services      |                 |
| General office clerks 2                  | Agility                   |                 |
| Information and communications tec 2     | Agricultural science      |                 |
| Electrical equipment installers and re 2 | Agriculture               |                 |
| Business services and administration 2   | Air decontamination       |                 |
| Machinery mechanics and repairers 2      | Aircraft electronics      |                 |
| Sheet and structural metal workers, 2    | Aircraft technology       |                 |
| Manufacturing labourers 2                | Airport management        |                 |
| Electrotechnology engineers 2            | Ambulance care (nursing)  |                 |
| Mining, manufacturing and construct 2    | Apache web server         |                 |
| Mobile plant operators 2                 | APL (programming langua   |                 |
| Architects, planners, surveyors and d    | Applications programming  |                 |
| Business services agents 2               | Applied mathematics (tea  |                 |
| Transport and storage labourers 2        | Applied statistics        |                 |
| Other clerical support workers 2         | Architectural models      |                 |
| Building finishers and related trades 2  | Archiving                 |                 |
| Blacksmiths, toolmakers and related      | Arts and crafts/hobby ma  |                 |
| Financial and mathematical associat 2    | Asbestos decontamination  |                 |
| Sales, marketing and development m 2     | Assembly (wood)           |                 |
| Building frame and related trades wo 2   | Assembly of parts         |                 |
| Database and network professionals 2     | Asylum and immigration I  |                 |
| Other sales workers 2                    | Attitudes                 |                 |
| Secretaries (general) 2                  | Audiovisual engineering   |                 |
| Mining and construction labourers 2      | , Auditing                | v               |

#### Workforce Migration



#### Statistical panel members

# Members of the <u>statistical</u> panel of evaluators of the European Big Data Hackathon

| Dario Buono        | Dario.Buono@ec.europa.eu   | Yes Yes  |
|--------------------|--|--|
| Christine Kormann  | Christine.Kormann@ec.europa.eu   | Yes Yes  |
| Jean-Louis Mercy   | Jean-Louis.Mercy@ec.europa.eu  | Yes Yes  |
| Aldo Laudi         | Aldo.Laudi@ec.europa.eu  | Yes Yes  |
| Geir Ottestad      | Geir.Ottestad@ec.europa.eu   | Yes Yes  |
| Pascaline Descy    | Pascaline.Descy@cedefop.europa.eu  | Yes Yes  |
| Tanya Chetcuti     | Tanya.Chetcuti@ec.europa.eu  | Yes Yes  |
| Stefano Abruzzini  | Stefano.Abruzzini@ec.europa.eu   | Yes Yes  |
| Silvia Biffignandi | silvia.biffignandi@ <u>unibg.it</u>  | Yes Yes  |
| Tobias Blanke      | tobias.blanke@ <u>kcl.ac.uk</u>  | Yes Yes  |
|                    | Dario Buono<br>Christine Kormann<br>Jean-Louis Mercy<br>Aldo Laudi<br>Geir Ottestad<br>Pascaline Descy<br>Tanya Chetcuti<br>Stefano Abruzzini<br>Silvia Biffignandi<br>Tobias Blanke | Dario BuonoDario.Buono@ec.europa.euChristine KormannChristine.Kormann@ec.europa.euJean-Louis MercyJean-Louis.Mercy@ec.europa.euAldo LaudiAldo.Laudi@ec.europa.euGeir OttestadGeir.Ottestad@ec.europa.euPascaline DescyPascaline.Descy@cedefop.europa.euTanya ChetcutiTanya.Chetcuti@ec.europa.euStefano AbruzziniStefano.Abruzzini@ec.europa.eusilvia Biffignandisilvia.biffignandi@unibg.ittobias Blanketobias.blanke@kcl.ac.uk |

#### Industry panel members

# Members of the <u>industry</u> panel of evaluators of the European Big Data Hackathon

| Oracle    | Mats Stellwall           | mats.stellwall@oracle.com                          | Yes Yes |
|-----------|--------------------------|--|---------|
| SAS       | Michel Philippens        | michel.philippens@sas.com                          | Yes     |
| Amazon    | Yotam Yarde              | yotamy@amazon.com                                  | Yes Yes |
| Microsoft | Frank Callewaert         | frankcal@microsoft.com                             | Yes Yes |
| SAP       | Gregory Neuveglise       | gregory.neuveglise@sap.com<br>gneuveglise@yahoo.fr | Yes Yes |
| Everis    | Sonia Taborcia Hernández | Sonia.Taborcia.Hernandez@everis.com                | Yes Yes |
| IBM       | Herman Nielens           | herman.nielens@be.ibm.com                          | Yes Yes |
| QLIK      | Jurgen Hamelrijckx       | Jurgen.Hamelrijckx@qlik.com                        | Yes Yes |
| AGILOS    | Edwin Willems            | edwin.willems@agilos.com                           | Yes Yes |
| Accenture | Thibaut Roisin           | thibaut.roisin@accenture.com                       | Yes Yes |

### EU BD HACK / PROTOTYPE CREDENTIALS

Live Prototype

# url: eubdhack.in2data.eu

## username: eubdhack@in2data.eu

## password: **EUBDHACK2017**

### Q&A