A Work Project, presented as part of the requirements for the Award of a Master Degree in Management from the NOVA – School of Business and Economics

# HUMAN-MACHINE SYSTEMS VS. THE UNEMPLOYMENT SPELL:

## HOW IEFP EMBRACED DATA-DRIVEN DECISION MAKING WITH PROFILING

**THOMAS DAUM**

**31758**

A Project carried out on the Master in Management Program, under the supervision of:

Leid Zejnilovic PhD

Thursday, January 3, 2019

**Abstract**

Data-driven decision making and well-developed analytical capabilities are generally perceived as fundamental for being a competitive organization nowadays. Nevertheless, especially publicly-led organizations show little agility towards technical advancement and face difficulties in developing necessary capabilities. The following case demonstrates how the Portuguese national body for employment and professional training, IEFP, engaged in a data-driven “profiling” model to combat long-term unemployment (LTU). The case walks the reader through the whole project-lifecycle, starting with IEFP´s previous touchpoints with data science over modeling and implementation of profiling, data curation, until managerial challenges which occurred along the way. The study reveals difficulties of a public organization linked to the usage of data-science and encourages students to look for ways on how to overcome those problems and push the progress forward.

**Keywords**: IEFP, Long-Term Unemployment, profiling, data-driven decision-making

**List of abbreviations**

LTU – Long-Term Unemployment

IEFP – Instituto de Emprego e Formacao Profissional

PES – Public Employment Service

KPI – Key Performance Indicator

GDPR – General Data Protection Regulation

ILO – International Labour Organization

**Human-Machine Systems vs. the Unemployment Spell:**

How IEFP embraced data-driven decision making with profiling

"*Long-term unemployment remains at very high levels in all countries that experienced a sovereign debt crisis. […] In most of these countries long-term unemployment affects the majority of job seekers.*"

Bertelsmann Stiftung

*"IEFP is facing major challenges with their effort in data-science. Several areas within the organization need refreshment. Counselors in the front offices do not want to lose their power of dealing directly with the job seekers; they show resistance against technical tools."*

Santana, IEFP

*"Clearly, profiling is needed to categorize jobseekers so that scarce resources are allocated in the most efficient way and so that they best serve the hard-to-place job seekers."*

European Commission

In March 2018, Cristina Faro, director of services of studies, planning and management control of the IEFP, was on her way to the weekly meeting with Cristina Taveira, team manager for statistics and responsible for the roll-out of IEFP´s profiling model – a data-driven approach to combat LTU. Today´s agenda assured a vivid discussion about the course of the profiling model. Already being in action for seven years, the general tendency within the organization was to carry on with the model. It enabled early intervention to prevent people from sliding into LTU and came with several other benefits, such as a detailed client segmentation due to its rich analysis of data. However, both were aware of the increasing internal doubts, especially coming from the counselors in the front offices who executed the profiling. Neglecting human-reasoned decisions through "blind algorithms", missing technological capabilities, and an outdated model had been among upcoming allegations. To discuss the pain points, a priority meeting, involving the board of directors and top-management of IEFP had been scheduled for the end of the month. For Taveira and Faro, the questions they needed to answer were evident: How could they improve the data-driven decision making and which were the necessary steps to make the profiling model more impactful? The deployment of the model reached a critical stage by then and only solid arguments would assure its continuation.

### Long-Term Unemployment[1]: A serious problem for society

Since 2009, the economic crises had immensely affected the EU, leading to historically high levels of unemployment and reaching a peak of 11%, on average, in 2013 (**Exhibit 1**). The impact of its consequences differed substantially within the European countries, with some states being able to better recover than others. For those tumbling countries, including Portugal, the unemployment spells were persistent and led to extraordinary high LTU rates (averagely 42% in 2017) (**Exhibit 2**). Consequences of LTU were linked to high costs not only for the individuals affected but also for the society and economy at large. Those costs implied poverty, increasing inequality, dysfunction of the economic system and a decrease in human capital. With around 50% being unemployed for longer than 12 months, Portugal ranged among the top ranks (6th position) of EU countries with respect to the highest percentage of LTU in 2017 (**Exhibit 2**). Thus, the government needed to focus on assessing LTU appropriately. The development of adequate activation strategies through public employment services (PES), especially through IEFP, represented a crucial measure to trigger the decline of LTU.

### The IEFP with the mission to fight unemployment

The institute for employment and vocational training, IEFP, represented the national body for employment and professional training. Its mission was to promote quality job creation and combat unemployment by implementing active employment policies, including vocational training through its career counselors.

[1] Per OECD definition, long-term unemployment means being unemployed for 12 or more consecutive months

The first ever dealing with unemployment by a public institution in Portugal took place in 1931. An unemployment fund was established to deal with the consequences of the economic crises at that time. From 1964 onwards first vocational trainings and national training centers were established to manage the adjustment between supply and demand for employment. These employment centers had continuously been expanded to ensure geographical coverage. As current training models no longer responded to the diversified needs of the labour market in the early 1970s, the IEFP came into play in 1979. Provided with financial and administrative autonomy by the Portuguese state, the IEFP was created with the objective to improve employment measures. IEFP´s goal was to better integrate employment and vocational training policies, considering the increasing diversity of the national territories: North, Center, Lisbon/ Tagus Valley, Alentejo, and Algarve. Since 2012, IEFP had been operating under the supervision of the Ministry of Labour and Social Solidarity with a flexible and autonomous organizational structure (**Exhibit 3**). The organization counted around 3.000 employees including central services, its five regional delegations, 30 employment and vocational training centers, 23 employment centers as well as one vocational training and rehabilitation center (**Exhibit 4**). Between 2007 and 2017 around 3.3 million citizens were registered in the IEFP system, of which 34% could be defined as LTU.

### Focus on clients and innovation

All actions of IEFP aimed at improving the lives of the unemployed population. This was being achieved by developing adequate action plans for the unemployed, once they were registered in the system (**Exhibit 5**). Maria José, Senior Technician, emphasized the focus on their clients, thereby keeping in mind IEFP´s social responsibility: "All our processes need to follow the guidelines of social responsibility. Dealing with the lives of people is a very sensitive topic and we are aware of that." For instance, a social responsibility department was installed and one of its tasks was to make sure that all actions were aligned and methodologies to combat unemployment were under permanent revision.

Besides the alignment to social responsibility, the organization also aimed for continuous adjustment to the needs of the market and improvement of efficiency. In this regard, innovation played a key role, as Cristina Faro underlined: “There is a lot of raw data and know-how within our organization which can be challenged in terms of innovation.” It was her and Maria José who served as “innovation accelerators” with the aim to trigger the whole IEFP. However, due to its heavy organizational structure and the fact that innovation was prompted mostly from external sources, they were facing several obstacles. There was no internal innovation drive within the organization due to conservative management philosophies which slowed down the adoption of new technologies.

**The course of data-driven management**

Since its foundation, all of IEFP´s business activities were based on data which the organization either accumulated by itself or received from external sources: “Data is a MUST and serves as a core to be able to manage the whole organization”, said Cristina Faro. In the early days, the collection of data mainly served evidence reasons. Nevertheless, IEFP had been one of the first public agencies in Portugal which developed key performance indicators (KPIs) based on their data to monitor business performances. Moreover, data was applied to better understand the employment market and act upon it, which remained a consistent driver until today.

The crucial problem for IEFP was to keep up with technology. “When we talk about data, it is not the amount which has changed but the technological tools and knowledge we developed over time to manage and extract information from it”, underlined Carlos Santana who had seen both sides of the coin as a counselor (front office) and statistician (back office). Besides the technological development, legal restrictions had always been a constraint for IEFP´s management of data, for instance with regard to transparency and access to client data. The latest example represented the General Data Protection Regulation (GDPR) by the European Union which the IEFP had to comply with in terms of client transparency. Legal restrictions also determined the course of profiling. Portuguese law required that people at high risk of becoming LTU had to be provided with more resources than the ones who faced lower risk. Eventually, profiling seemed to be the most suitable approach to meet those requirements.

## Profiling and bumpy roads to proof an impact

### *A call for a more effective approach*

Chronical unemployment – people facing recurrent spells of unemployment – and underemployment – prevalent in economies where a large proportion of part-time workers seeks for more working hours – had been detected as a major cause of the very high LTU rates in Portugal and other Southern European countries such as Greece, Spain, and Italy. LTU rates in those countries went above 10% in 2013. A poor structure of employment policies in those European economies resulted in a deficient allocation of resources, in particular in a limited supply of job offers for those seeking to be employed.

To improve the efficiency and accuracy of active employment policies, profiling was implemented. As such, profiles were created by comparing characteristics of individuals newly unemployed to those of the LTU. The purpose was twofold: *i)* to better identify those being at most risk of becoming LTU, and *ii)* to allocate resources adequately. PES in France and Germany, for instance, followed a coherent and integrated strategy fully based on the outcomes of profiling. Others, such as Sweden and Ireland, did not entirely rely on profiling. They used administrative data of the profiles to support their decision-making processes.

***Profiling at IEFP: From risk score to actions***

Already in 2011, a task force at IEFP with Maria José as project leader had been instructed by the Ministry of Labour to develop a profiling model as intervention methodology with the unemployed. The task force deployed a few logistic regression models and thereafter the results were segmented into three different risk categories: *low, medium & high risk* (**Exhibit 6**).

The segmentation into profiles helped counselors working in front offices to determine suitable courses of action for the unemployed to either regain employment in a similar profession or change career paths. As Cristina Faro underlined, "the model was very important and streamlined for our counselors because they could immediately infer actions and allocate them to the potentially high risk unemployed." For instance, if the model determined that a registrant was only at low risk, the counselor would have chosen a job interview as the best course of action. In the case of high risk, a more careful analysis of the case with follow-up counseling services was necessary. For Carlos Santana especially the differentiated segmentation feature represented a major benefit: "The profiling model enabled a sophisticated analysis of the LTU by providing us with a detailed overview of their characteristics, which would eventually minimize the probability of them becoming LTU."

Nevertheless, the IEFP profiling team had several strong concerns. The model was calculating risk only at the time of first registration at IEFP. Would it consider to what extent the population changed since 2011 then? And would the same model apply to the new reality with regard to the labour market?

***What makes a good data science project?***

The task force led by Maria José brought in the internal analytical capabilities in terms of developing statistical models, while Cristina Taveira was responsible for determining the project scope and reporting to the top management. "Although the team size was small and the project demanded a lot of commitment from everyone involved, we were convinced that we had the right data, analytical capabilities, and organizational maturity to face the project", said Cristina Taveira.

When the project was executed, an action-driven framework consisting of four basic steps had been applied in order to scope the project adequately: Define goals, inform actions, receive and work with data, and eventually an analytical approach (**Exhibit 7**). The goal of the project was to decrease the LTU rate as much as possible. More specifically, to better identify people with a high risk of becoming LTU and to allocate resources adequately. Actions were informed by defining individuals with a certain LTU risk score who would then receive a combination of measures, such as interviews or vocational training. Based on the data (*see data part p. 8f.*), the IEFP could analyze which combinations of actions worked best for each individual.

***Profiling against LTU: A winning approach?***

On a European scale, most countries made use of profiling as part of their active employment policies to combat unemployment (**Exhibit 8**). The benefits of a detailed client segmentation and an early identification seemed to be key elements for using profiling and eventually improving the effectiveness of their activation measures. Yet, PES adopted profiling models at varying levels of sophistication, using exclusively or combining *counselor-based profiling* (counselors are exclusively responsible for the evaluation of an individual´s employability), *rules-based models* (allocation based on personal client characteristics, such as age and gender), *statistical* (decision-making solely based on statistical analysis of data) and *data-assisted profiling*.

The most used methodology had been *data-assisted profiling* where counselors retained their influencing role in client segmentation but were provided with quantitative data which they could use for the purpose of diagnosis and informing actions. Thus, the more data had been available and could be added to the databases of PES, the better would be the opportunity of predicting LTU among individuals registered at the organizations. Nevertheless, the automation of results generated by "blind" algorithms was perceived critically among PES. The right balance between "man and machine" in profiling still had to be determined.

## IEFP´s data about the unemployed: Learning the context and collaboration modes

### *Data composition and technical methodology*

"All the data we had access to was fundamental to ensure the profiling of our clients and the functioning of our services in general", said Cristina Taveira. To guarantee a high-quality composition of data, the IEFP relied on years of collected transactional data about individual clients, including socio-demographics (age, gender, location, etc.), and their interactions with the system, such as job interviews, trainings, attended meetings and job offers made by companies. Moreover, professional characteristics, such as area of expertise, desired careers and years of experience were collected by the IEFP. Because the organization had to deal with very large amounts of information, they separated the received data into five tables based on the type of each individuals´ interaction with the organization: *Requests*, *Summoned*, *Interventions*, *Presented* and *Job offer* (**Exhibit 9**). Only *Requests* tables for instance, which documented the beginning and end of a registrant relationship with the IEFP, contained 24,695,518 records for 3,338,363 unique individuals.

In terms of technical methodology, IEFP´s profiling model, when being rolled out, consisted of two logistic regression models (one for each gender) and predicted probabilities of being LTU by modeling three levels of risk (low, medium and high). Later, in a second iteration phase individuals were divided into four categories: men looking for 1.1) first jobs and 1.2.) new jobs, and women looking for 2.1) first jobs and 2.2) new jobs. The profiling model was run once per registrant at their first touchpoint with IEFP, neglecting transactional features, such as future intervention patterns. When exiting the system and re-entering at a later stage, the counselors in the front offices had to manually update the individual´s risk score. Trainings and workshops would not affect the risk score of a registrant unless the counselor updated it explicitly.

***Data sharing***

Not only with regard to profiling, the sharing of data was vital for the functioning of IEFP. “What we should not forget about our data is that we hardly collected data by ourselves and relied heavily on external sources like the National Institute of Statistics or Ministry of Labour. This way we used the data to derive our actions and produce additional statistics”, underlined Cristina Faro.

Important players within IEFP´s eco-system were social securities, the European Union with several entities, government bodies and other PES. Social securities, for instance, shared valuable and legally protected insights, such as (un-)employment periods of individuals. The European Union with national, regional and local networks contributed with knowledge about their experience working directly with unemployed individuals. Moreover, universities were on the rise to become an important player due to their growing research activities and knowledge in data science. Those collaborations were essential because data sharing allowed IEFP to have a more detailed view of the job seekers and the market. Once the data had been shared, collected and updated into IEFP´s information system, it would enrich the accuracy of profiling and serve the top-management of IEFP to better take decisions about active measures of employment.

**Is Profiling the way to go?**

Getting closer to the upcoming priority meeting, the majority of IEFP´s management was still convinced to have introduced a beneficial approach with profiling. Nevertheless, several problems remained and new critical challenges occurred since its implementation in 2011, especially in terms of organizational maturity. In late March 2018, IEFP´s decision makers of departments involved in the process and the board of directors eventually met to discuss the future usage of profiling.

*Technological maturity* represented one major discussion topic. There were several doubts if the model was able to cope with atypical periods, such as economic crises because it was not robust to temporally changing conditions. IEFP´s model was static and could not dynamically update profiles for each time an interaction with the system took place. Once registered in the system, the risk score would remain the same if not updated manually by the counselor.

*Evaluation* was another pain point, being criticized by Maria José: "How shall we assess the impact of the model if there is no way of evaluating it yet?" Indeed, there was no information in terms of effectiveness of measures (i.e. the impact of proposed actions to the unemployed) which had been taken based on profiling data. Consequently, it could not be determined which action had been the trigger for people exiting the system. The only aspect which had been evaluated, so far, was the accuracy of the calculation of profiles. With 70% accuracy, predictions were imprecise which could result in incorrect classifications of people being at high risk of becoming LTU.

It was the *project scoping* part which especially concerned the board of directors. Referring to the small task force responsible for the execution, they highly doubted that analytical capabilities and in-house resources would suffice. As such, they suggested to reach out to experienced people with strong skills in both, data science and project scoping when applying an action-driven framework (**Exhibit 7**). For them, a difficulty was already defining a clear analytical goal. Namely, profiling as means for fighting LTU had been vaguely defined. What was IEFP planning to maximize/ minimize? Were there any constraints (budget, resources, etc.)? By focusing on a more concrete goal, such as increasing the probability to respond to a critical mass of the population, the board members believed that actions could be informed more precisely, consequently leading to a better allocation of resources.

The majority of participants agreed that the profiling features of early intervention followed by the allocation of suitable measures would lead to cost savings and a reduction of the caseload for counselors. However, the problem for them lay in HR *resources*. Appropriate services such as intensive counseling and training courses needed to be available to support the profiled jobseeker. The job market was changing more than ever with a growing diversity in society and IEFP as well as other PES did not seem to have a sufficient number of staff to keep up with this change. As a result, this would hinder the delivery of tailored services identified through profiling.

A critical topic which popped up repetitively during the meeting was the management of *organizational communication* and missing *technological knowledge of the counselors*. Due to the lack of communication and missing education, IEFP´s counselors were skeptical about the application of profiling. Not being familiar with the technology, counselors perceived the automation of results through profiling models as a replacement of reasoned decisions by humans. They were sometimes expecting different profiles from the model due to their prior familiarity with the capabilities and paths of a registrant, leading to a decision-conflict. For Cristina Faro, this represented a big challenge: "Several areas within the organization need refreshment of their technological knowledge to see what is possible with data science. Our counselors play a vital role in the whole profiling process. If they received explicit instructions in the first place and if there was continuous communication between the different departments, they would feel more confident in using the model and leverage from the opportunities of data science."

*Social Responsibility* and behaving *ethically* correct had been permanent issues in IEFP´s priority meetings since dealing with the lives of people represented a highly sensitive matter. Those participants of the meeting who criticized the profiling model perceived it as discriminating. For them, the categorization of unemployed generated a social sorting in which some people were better than others, neglecting equality of treatment. Simultaneously, part of the IEFP management in favor of profiling claimed that discrimination was always present, no matter which methodology was in place: "Even if resources were allocated to a whole LTU population without segmentation, it would cause the problem that resources would be given to people which do not need them to the same extent as others" , said Cristina Taveira. This, in turn, would result in resource-based discrimination. For them, it was more a question of how to keep discrimination to a minimum. For instance, by looking at the different components of the data, such as its entry, processing, and analysis, potential biases (i.e. automation biases[2]) embedded in the data could be identified and approached in a streamlined way.

Another topic on the agenda was the involvement of *partnerships* which contributed to the model. IEFP did not produce innovation internally. It was very important for them to get teased by external sources due to their limited intrinsic motivation. Only with the help of partners offering cutting-edge technology related to the processing of data, and other synergies, such as huge datasets to build upon, the current profiling approach could be pushed forward and become sustainable.

While everyone agreed on the benefits of embracing partnerships for the sake of profiling, the last critical topic remained silent, namely IEFP´s *dependence on government bodies*, like the Ministry of Labour and Social Solidarity. The coherence to the governmental goals always had to be ensured. Being the government´s executive body, the IEFP was facing limitations when

[2] Tendency to heavily rely on automated systems which can lead to incorrect automated information disabling correct decisions

designing processes and pursuing their own goals. The dependence drastically influenced its policies so that IEFP could not use data in the best possible way to serve job seekers and companies. It was obvious for the management of IEFP that a clearer spin-off from political bodies would help them to be more independent in order to act efficiently. However, the feasibility of this matter could be questioned.

**What´s next?**

Although a lot of doubts had been expressed during the priority meeting, the participants came to the conclusion to carry on with the profiling model since the project was in a crucial phase and still came with a lot of unexploited potential. As one course of action and to carry out essential adjustments, thanks to their network with science and research institutions, the IEFP decided to start a collaboration with the DSSG – a program called Data Science for Social Good Europe created by Nova School of Business and Economics – to design an updated LTU risk prediction model. With a team consisting of a mixture of aspiring data scientists and experienced project managers, the IEFP management was convinced to cover some of the detected pain points.

In order to develop an updated model, the DSSG team first needed to eliminate errors and out of the scope cases in IEFP´s data. By gathering and curating datasets of the available and also additional macroeconomic indicators, the economic conditions of unemployed individuals could be represented better, leading to a better data infrastructure and eventually to more sophisticated individual´s profiles. Moreover, through the transformation of raw data tables into analytical tables Machine Learning algorithms could be run on the data. To understand their adequacy of predicting future scenarios, the team split the data into train sets, validation sets, and test sets according to time. The measurements taken by the DSSG team promised several technical advancements for the profiling approach: being robust against changing temporal conditions; generating risk scores dynamically with every interaction; providing an evaluation metric & easy-readable explanations about which features had an impact on an individual´s risk score; and offering region-specific models.

With the partnership in place, the immediate question popped up if this was the right direction to go for IEFP. Did they manage to cover most of the paint points which came up in the priority meeting? Which challenges still remained? How could it be ensured that counselors eventually let go from their skepticism about the model? How would the evaluation, in particular, critical success factors look like? Which measures had to be taken to ensure that the project was sustainable and had a long-term impact?

**Exhibit 1**: Euro Area and EU28 unemployment rates (%), in *Eurostat, July 2017*

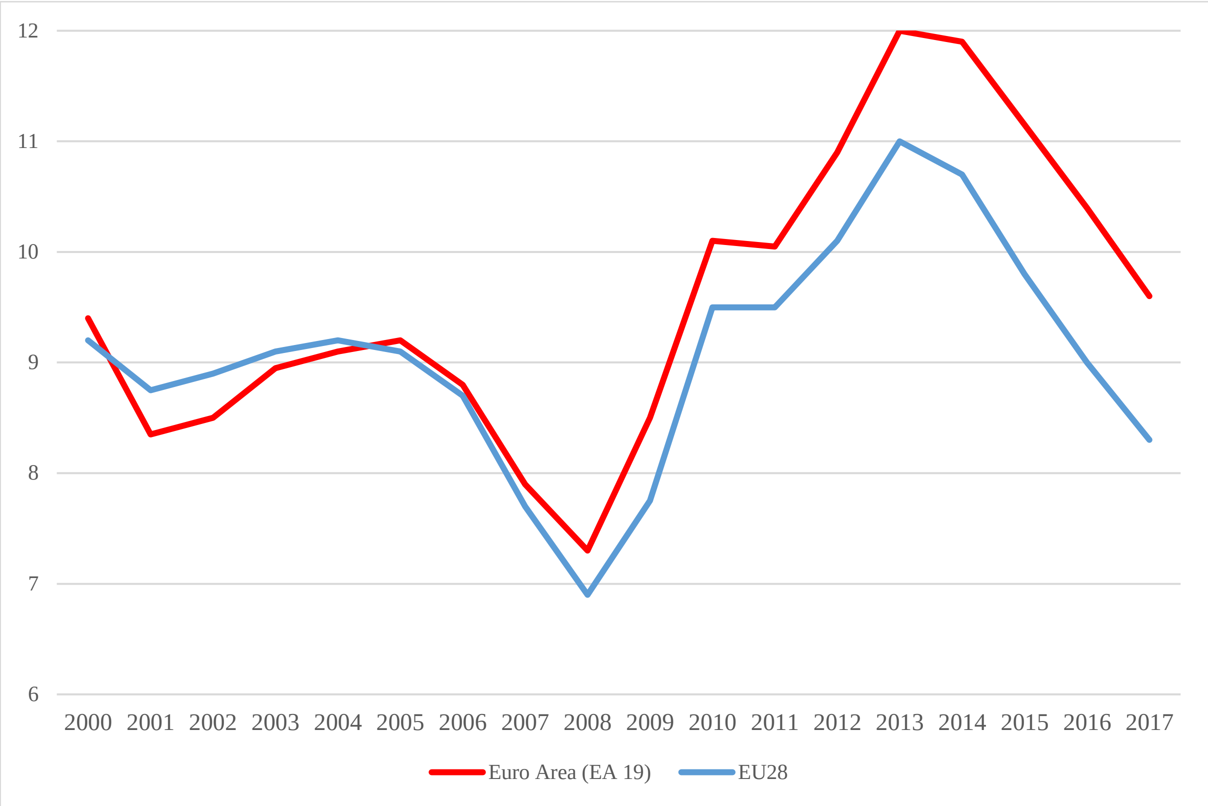



**Exhibit 2**: EU Long-Term Unemployment rate. % of unemployed in 2017, in *OECD Data*

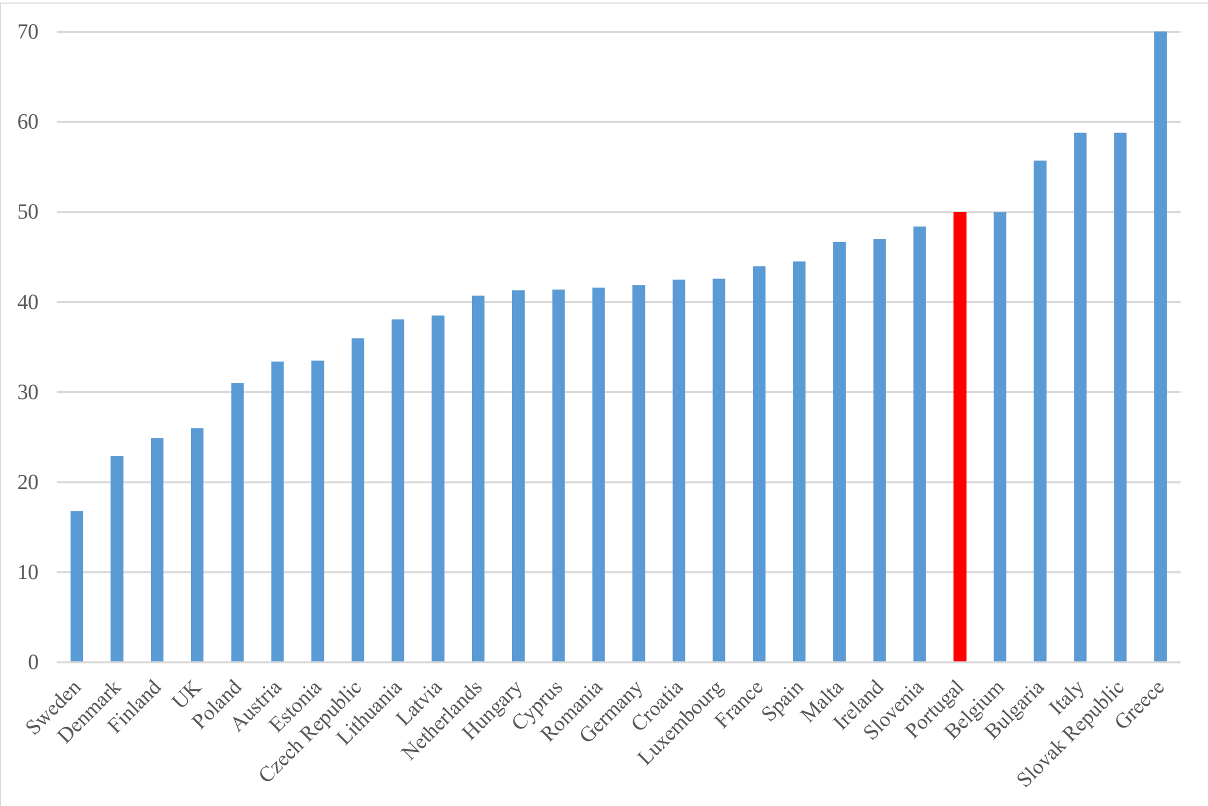


**Exhibit 3**: Timeline of important events throughout the history of IEFP, in *IEFP online/ History*

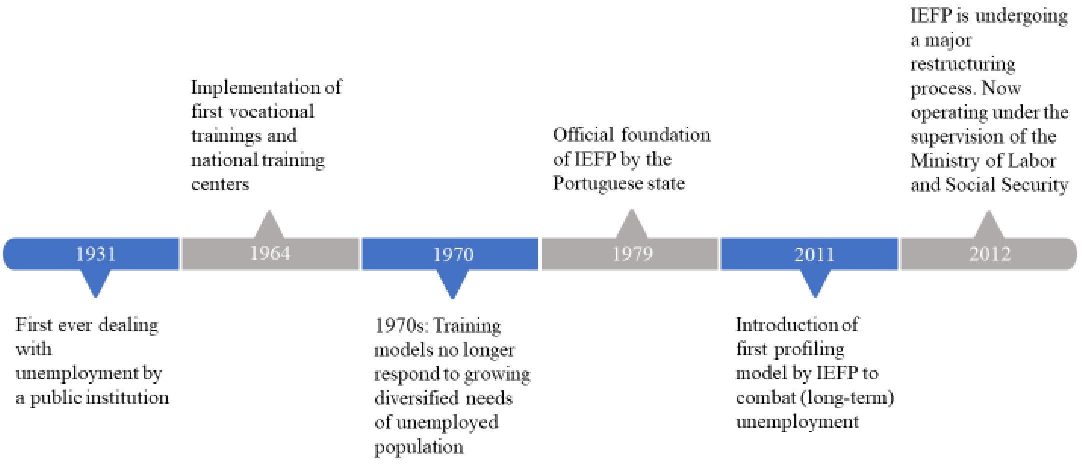



**Exhibit 4:** Five Regional divisions, in *IEFP online/ Structure*

**Exhibit 5:** How does IEFP work?, in *Predicting the risk of long-term unemployment in Continental Portugal, Final Poster (DSSG), August 2018*

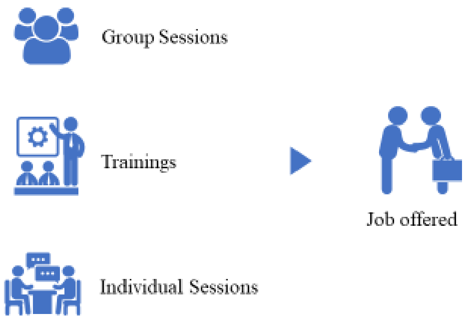



**Exhibit 6:** IEFP Profiling model: Predicting Long-term unemployment risk score, in *Cascais Data Science for Social Good Europe Summer Fellowship 2018*

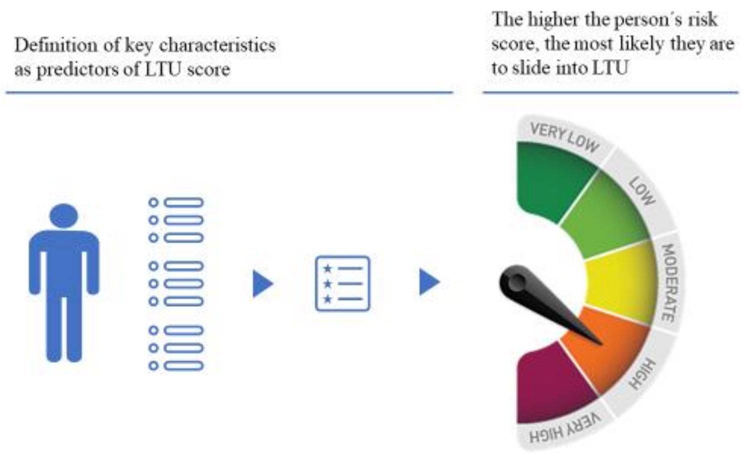


**Exhibit 7**: Action-driven framework (Project scoping), *in Scoping Data Science (for Social Good) Projects, 2016*


***Step 1: Goals** – Define the goal(s) of the project*
***Step 2: Actions** – What actions/interventions do you have that this project will inform?*
***Step 3: Data** – What data do you have access to internally? What data do you need? What can you augment from external and/or public sources?*
***Step 4: Analysis** – What analysis needs to be done? Does it involve description, detection, prediction, or behavior change? How will the analysis be validated?*

**Exhibit 8**: Usage of profiling in selected OECD countries, in *Tackling Long-term Unemployment through Risk Profiling and Outreach, May 2018*

| Use of profiling | Focus | Description |
|---|---|---|
| Diagnostics | Client Segmentation | Profiling differentiates clients based on unemployment risk diagnostics |
| Targeting | Action plans and allocaiton to employment programs | Prior diagnostic profiling results help the caseworker to define client needs and draft a mutually agreed employment acion plan |
| Resource Allocation | Budget planning and policy controls | Profiling can help PES management plan fiscal resources based on severity of client profile, and inform PES management in aligning policy with resources based on setting and altering cut-off risk thresholds |
| | ALMP contracts | Profiling can help PES caseworkers to contract out employment programs to private sector providers |
| Additional Applications | Enrichment of labour market statistics and aggregate skills profiling | Profiling of jobseekers gathers information that can enrich labour market statistics whiich can be useful for understanding dynamic changes over time. Profiling can support efforts to conduct macro-level skills needs assessmnets based on aggregate skills profiling |
| | Enhancement of job matching | Client profiles can support PES caseworkers to better match job seekers with available vacancies through job crawling mechanisms |

*ALMP – Active Labour Market Policies*
*PES – Public Employment Services*

**Exhibit 9**: Data composition in form of five different transactional tables, in *Predicting the risk of long-term unemployment in Continental Portugal, Final Report (DSSG), August 2018*

| Type of table | Description |
|---|---|
| Requests | Individuals that request job-seeking support from IEFP (with information on entry, update and cancelation of requests) |
| Summoned | Individuals that are requested to present themselves at their local IEFP centers (or to get in contact with them); these communication attempts can be made either physically or electronically (e-mail, letter, phone call, etc.) |
| Interventions | Individuals that were assigned to or attended an intervention facilitated by IEFP |
| Presented | Individuals that attended a job interview |
| Job Offer | List of job offers |

**Case References**


Data Science for Social Good (DSSG) Europe (2018). *Data Science for Social Good Summer Fellowship.* Retrieved from https://dssg.uchicago.edu/europe

Data Science for Social Good (DSSG) Europe (August 2018). – Final Report: *Predicting the risk of long-term unemployment in Continental Portugal.*

Faro, C. (September 27th, 2018). *Personal Interview*, IEFP, Braço de Prata, Portugal

Ghani, R. (October 216). *Scoping Data Science (for Social Good) Projects.* Retrieved from https://dssg.uchicago.edu/2016/10/27/scoping-data-science-for-social-good-projects/

IEFP (n.d.). Retrieved from: https://www.iefp.pt/

IEFP Quality Letter (April, 2018). *IEFP Carta da Qualidade*

Taveira, C., José, M., Santana, C. (September 14th, 2018). *Personal Interview*, IEFP, Braço de Prata, Portugal

## Teaching Note: Human-Machine Systems vs. the Unemployment Spell

### Case Synopsis

With 465,000 citizens (6,8%) being unemployed in 2018, Portugal shows one of the highest unemployment rates within the EU (**Appendix 1**). Half of those unemployed citizens are LTU. A status that is given to those without a job for longer than 12 months. LTU has heavy consequences for unemployed individuals (i.e. negative psychological conditions) and the society (i.e. dysfunctional economy). Portugal´s public employment agency, IEFP, is constantly aiming to combat unemployment through active employment policies. Fighting unemployment since 1979, the IEFP has become the most important player in national unemployment matters. By having introduced a profiling model as main methodology to fight unemployment, the IEFP has established data-driven decision-making in 2011. However, due to the rapid technological advancement and rising internal doubts, the management needed to identify pain points and potential improvements of the current model. Amongst others, technological maturity, social responsibility and project scoping had been identified as major issues to overcome.

The first part of the case outlines a historical overview and background of the IEFP. It also introduces its´ first touchpoints with data and walks one through the introduction of profiling. The second part focuses on the different elements of profiling including technological methodology, data sharing and curation. Eventually, major difficulties of IEFP as a public entity associated with the usage of data science are revealed, which provides the reader with sufficient impulses to look for ways how to overcome those challenges and benefit from data science.

### Learning objectives

The IEFP case provides students with an overview of how a publicly-led entity engages in data-driven decision making to combat unemployment. Specific learning objectives include: ***(i)*** To examine the rationale of introducing data science and adopting analytical approaches inside a public entity; ***(ii)*** To analyze the managerial challenges of integrating and sustaining data science in a public entity; ***(iii)*** To understand an exemplary data-driven approach, profiling, used to combat unemployment.

The IEFP case can be taught in intermediate Master level courses and covers various areas, including Project Scoping, Organizational Change Management, and Business Ethics. Furthermore, the case can serve as complementary discussion material for courses which have a strong emphasis on data science, such as Data Curation and Visualization.

**Case Analysis**

To ensure an informed discussion on IEFP´s profiling model, it is crucial to understand the contextual factors: The lecturer could perform the case analysis by asking which stakeholders are involved in profiling and what are the interdependencies between them. Looking at each stakeholder´s point of view also helps the students to get an overview of the different key areas which the case is addressing. Four stakeholders can be identified in the case: ***(i)*** *Government* – gives the authority to the IEFP to execute actions; ***(ii)*** *IEFP: Top-management* – decides upon the best way to achieve the objectives defined by the government; ***(iii)*** *IEFP: Counselors* – execute the profiling and resource allocation, and translate the managerial decisions and processes to the clients; ***(iv)*** *Unemployed population* – the most affected stakeholder by profiling, gives mandate to the government to orchestrate problem-solving.

**The Government´s Point of View**

This section examines the role of the government and the extent to which it influences profiling. Four different dimensions emerge when analyzing the stakeholder: *responsibilities (initiator), international benchmarking*, *legal restrictions,* and *ethics*. Before analyzing each dimension, students may firstly explore the relationship between the government and its citizens. The government is continuously exposed to socio-economic pressure, and the management of unemployment can be seen as an indicator of social unrest (Nye et al., 1997). It is responsible for the stability of the economy and the wellbeing of its citizens. For this reason, the government invests in re-integration measures to combat unemployment, from insurance to intensive counseling through PES (ILO, 2014). Given its role as an *initiator*, the government gives mandate for the execution of the measures to PES. Students should infer from the case that it is not the IEFP, but the government that decides for profiling. This decision took place mostly based on *international benchmarking*. There are two common approaches used by national bodies to combat LTU: *i)* implementing early intervention initiatives to avoid that people slip into LTU; and *ii)* proactively assisting them to foster re-integration processes (Scopetta and Buckenleib, 2018). In this regard, efficiency and timing play a crucial role. For instance, insurance benefits have little effects on job-finding in weak labour markets and can cause longer unemployment periods (Faber and Valleta, 2014). In terms of timing, early intervention reduces the risk of people becoming LTU. Since the early 1990s, governments explored profiling approaches, which seemed to respond best to timing issues and efficiency (Rosholm et al., 2004). Despite setbacks (i.e. accuracy), profiling is nowadays used in most countries to combat unemployment (Rudolph and Konle-Seidl, 2005). A further dimension to consider are *legal restrictions*. Portuguese law[3] requires that people at high risk of becoming LTU must be provided with more resources than the ones who face lower risk. Profiling corresponds to these legal restrictions by generating risk categories for each registrant. These aspects, coupled with technological advancements such as Machine Learning algorithms, more efficient data infrastructure, and consequently, a promising way to allocate scarce resources, led the Portuguese government to profiling (ILO, 2014). Students should also consider *ethical* conflicts for the government. Should the government give priority in resource allocation to the ones

[3] Legal basis: "Within the scope of the Commitment for Growth, Competitiveness and Employment, signed in January 2012, and of the Program to Relaunch the Public Service of Employment (Resolution of the Council of Ministers, nº 20/2012, 9th of March)"

which can get out faster from unemployment? Or should the worst affected be elevated to level the playing field in competing for a job? There is also the acceptance of data-driven decision making – What is the right balance between man and machine (Eichhorst et al., 2015)?

**IEFP´s Point of View**

The following section analyses the stakeholder and entity that represents the core of the case study: the IEFP. A division into *top-management* on the one side and *counselors* on the other is recommendable, as it allows capturing the tensions that may occur within the organization.

*Top-Management*

The IEFP acts as the executive body of the government. It receives defined goals by the government and decides about how to achieve those. In the context of profiling, the top-management had to find a way to design an approach which decreases the LTU rate as much as possible. The aspects below contribute to the better understanding of IEFP´s respective actions.

*Importance of Data Science*: In terms of developing the profiling initiative, the lecturer may refer to data science and debate about *i)* its relevance and *ii)* added value for IEFP. Looking at the purpose of data science, it intends to extract knowledge from data and inform actions. Data science covers contextual understanding, data collection, analysis, and retrieving insights to enable practical implementation. Machine Learning algorithms and other forms of advanced analytics transform data into valuable knowledge to make it more meaningful (Das, 2016). Having introduced data science, the teacher may want to bring the discussion into the context of combating unemployment. Data has been prevalent in unemployment organizations since the 1980s. But, cost of computation and scarcity of Machine Learning algorithms were limiting the use of data (Guerrero and Lopez, 2016). It is of great interest for the top-management to make the profiling model as accurate and efficient as possible. Reflecting on the many types of data the IEFP collects (i.e. data about entire unemployed populations), it can leverage on the combination of this data with advances in data science. As such, the data would be more meaningful, actions could be better informed, and the top-management could achieve the defined goals by the government.

*Technological maturity*: The IEFP has rich databases with access to a lot of data. However, its technological maturity to benefit from this data is questionable. Technological maturity can be seen as a limiting factor for the top-management for several reasons. With one technology disrupting the other in short cycles in the context of data science, the IEFP has to permanently create and re-create structures to be able to cope with them. PES like IEFP face heavy hurdles regarding the adoption of data science. Conservative management philosophies and legacy systems refer to a sector which is historically slow to adopt new technologies. Traditional structures hinder internal innovation drive. Also, concerns about data security and trust in technology play a crucial role. There is no data trust among the institutions to exchange data securely (Nagy, 2016). Once having decided for a data-driven approach, a significant degree of robust infrastructure to handle the information and profound analytical capabilities are necessary to extract meaningful information from the data. Moreover, the collection and delivery of information must correspond to data governance rules and be aligned with every business stakeholder involved (Kotarba, 2018). Considering the prevailing limitations, setting data-driven decision-making as a business priority represents a critical issue for the top-management of IEFP.

*Decision for Profiling approach*: A closer look at different types of profiling may provide the students with sufficient information to understand the reason why the top-management decided for its particular profiling model. According to Loxha and Morgandi (2014), there are four approaches, which can be used exclusively or in combination, based on data availability and ability to process this data (**Appendix 3**). *Caseworker-based profiling*: Counselors are responsible for the probability of job seekers´ employment prospects. The focus lies on qualitative methods, such as interviews. It is linked to high HR resources and limited access to data. *Rules-based profiling*: is applied with time-based (length of unemployment) and demographic (age) segmentation, and is cost-efficient as only basic information is needed. It lacks precision and can cause deadweight effects. *Statistical profiling*: is done based on statistical analysis of data. The feasibility depends on the availability of data as the data is used to predict the probability of registrants finding a job. *Data-assisted profiling*: is a combination of *counselor* and *statistical profiling*, most spread in practice. Most of the decision-making power remains with the counselors as they get assigned a key role. Quantitative data serves as diagnostic support. It is the task of the top-management to decide for a certain profiling model. On the one hand, due to the data-reliant unemployment landscape and the advances in analytics, data-driven insights seem to be indispensable. Meaningful information like underlying patterns about populations can only be extracted adequately by leveraging on data science. On the other, the interests of counselors as executors of profiling need to be preserved. Considering their key role in profiling (*to be explained below*), *data-assisted profiling* appears as logical consequence.

*Project Scoping*: The lecturer may refer to the top-management´s way of applying an action-driven framework when implementing profiling (i.e. Project Scoping course). In the case, the action-driven framework by Data Science for Social Good, was applied (DSSG, n.d.). Once the management determines the general objective, it can proceed with identifying actions and types of analysis that will inform the actions.[4] The case provides examples where the top-management is facing difficulties. *Goals* are critical as the subsequent steps are formed according to them. It is crucial to determine analytical goals to identify how they may influence overall goals. The initial goal of reducing LTU in Portugal is defined vaguely, without any quantitative objectives and neglecting constraints. This is a common issue and indicates a need for iterating scoping

[4] For reasons of simplification and to help students structure a project, the teacher could hand out a *Data Science Project Scoping Worksheet* (**Appendix 3**)

processes so analytical goals can be better defined (DSSG, 2018). Being more precise on goals would help the top-management to formulate the subsequent steps and eliminate uncertainties. A specific goal could be to increase the accuracy of predictions for X percent. Secondly, only *actionable* projects can create an impact. Thus, actions need to be specific and feasible. The top-management needs to determine resources required to achieve declared goals. There are both limitations in resources (staff) and political dependencies that limit streamlined actions. Once having determined goals and respective actions, the *data* has to be examined. Which data is accessible for the IEFP? Which data is needed to solve the problem efficiently? It is the duty of the top-management to ensure a robust infrastructure to handle the data. They can leverage on a rich data system, having collected data since 2007. However, the system is 20 years old and shows siloed databases with different rules for data collection and storage, making it difficult to extract and transform the data. Also, the IEFP neglects external macroeconomic data to better represent economic conditions. To inform actions, the IEFP has developed a predictive model to *analyze* individual´s LTU risk. Predictive modeling means predicting the probability of a category, for instance, which level of risk a registrant belongs to. Consequently, suitable courses of actions can be informed. Lastly, validating the analyses should also be part of a good project scoping. The top-management has a predictive model in place but struggles in evaluating actions as the right metric still has to be found (so far only by accuracy of the results).

*Ethics*: Ethics is a sensitive topic associated with the management of people. Every measure taken by the IEFP is perceived critical among its stakeholders. With its profiling model in practice and access to holistic data, the top-management is confronted with several types of biases. Those are embedded in the different components of the data. Potential biases occur in the entry of data, go along its processing and reach until the analyses and interpretation. Several kinds of cognitive decision-making biases can lead to irrational judgment: Focalism[5],

[5] Tendency to depend on a specific trait of information when taking decisions – cf. risk score (profiling)

Automation biases[6] and belief biases[7], just to name a few. Also, forms of indirect discrimination caused by a certain outcome of profiling can result in situations unfavorable for a person due to disability or age. It can be argued that if sensitive data is involved and if a lack of transparency exists, the risk of discrimination increases respectively (Niklas et. al, 2015). For the top-management, keeping biases and discrimination to a minimum is therefore highly relevant. To do so, they need to be aware of the sensitivity of the data, every stakeholder involved and their impact along the profiling process. Diminishing potential biases for one stakeholder means an increase for the other (i.e. man versus machine). Consequently, biases will always be present. By designing processes carefully while considering the sensitivity of data and the impact of each stakeholder, the top-management could, however, keep it to a minimum.

*Counselors*

Counselors, representing the other part of IEFP, assume a key role in profiling. They are the ones executing profiling and allocating resources to the unemployed, being responsible for the outcome. There is evidence among countries practicing profiling that the involvement of counselors in terms of developing, using, and interpreting the profiling model is vital to success (Scopetta and Buckenleib, 2018). At the same time, statistical profiling using advanced analytics receives wide acceptance within many PES. In turn, this leads to the dilemma of finding the right balance between man versus machine (Barnes et. al, 2015). In fact, profiling models based on advanced analytics and algorithms struggle with complex individual situations of clients, which leads to over-simplifications or mistakes. To overcome this shortcoming, human intervention is necessary. But granting counselors too much freedom could result in using their own judgments to an extent that processes can be manipulated (Niklas, 2015).

[6] Tendency to heavily rely on automated systems which can lead to incorrect automated information disabling correct decisions – cf. risk score (profiling)

[7] Effect where an individual evaluation of the logical strength of a statement is biased by the plausibility of the conclusion rather than how strongly the statements support the conclusion – cf. risk score (profiling)

Several doubts arise for counselors. Firstly, they do not want to compromise on their power and show resistance against the model. Secondly, they are missing technical knowledge due to a lack of internal communication. The implementation of advanced analytics, however, implies employing staff with the education needed when performing complex tasks in the field of data management (Nagy, 2016). Lastly, they do not trust the results of blind algorithms. There is no clear tendency yet to which extent counselors should be involved in profiling and what is the best way to educate them. Advanced analytics in form of Machine Learning cannot replace human assessment. Such tools only serve as support in decision-making and should never be approached as an equivalent to human decisions. To facilitate this transitional relationship and simplify the lives of counselors, a focus has to lie on communication and education. To enable effective organization-wide decision-making, communication and collaboration are necessary. Like this, existing silos, such as the counselors' prejudices can be overcome.

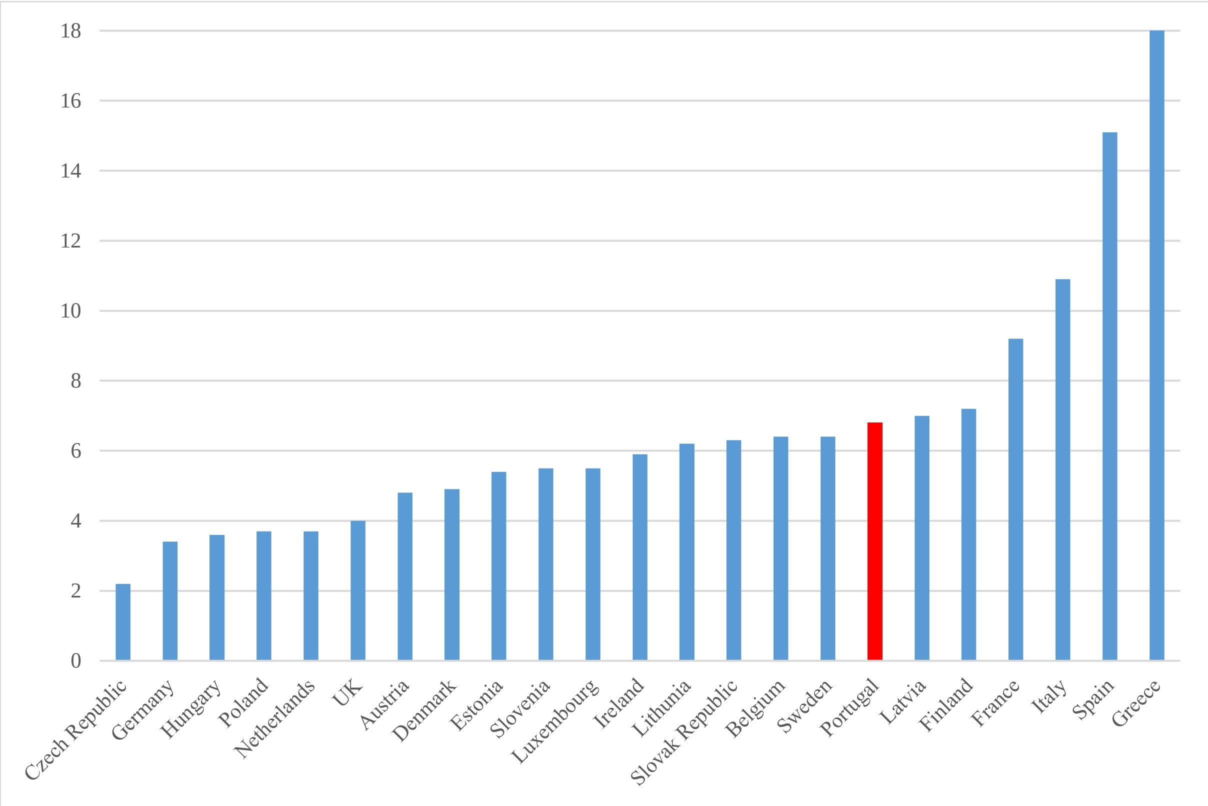
**The Unemployed Population´s Point of View**

The last section analyzes the most affected stakeholder: the unemployed population. Hindering their activation in the labour market leads to significant costs, such as developing psychological conditions which can result in low motivation to seek for new opportunities, depression, somatization and poverty (Kroft, et al., 2014). IEFP´s profiling model is perceived critical among the unemployed. What they require from the measures taken by PES is to be employed again. The question of how the re-integration works seems to be of secondary importance. However, whether to accept decisions from algorithms vs. humans also plays a crucial role. What does it take for the unemployed to turn their back on the expertise of counselors and follow automated results (and vice versa)? In this regard, *Transparency*, *interpretability*, *effectiveness*, and *ethics* are aspects to look at. In terms of *transparency*, unemployed individuals experience uncertainty in connection with the processing of data. They have limited access to information on which actions are taken in the course of profiling and how resources are allocated based on a certain profile. Criteria such as age, gender, etc. should be transparent for those involved. Transparency is accompanied with the *interpretability* of outcomes. With the usage of data science, the complexity of data is likely to rise as well. Human-readable explanations about which features have an impact on an individual´s risk score not only help the technicians but also the affected to better understand why they are allocated to a certain risk. Ways of *evaluating* profiling measures still have to be refined. The IEFP has not found the right metrics for its approach yet. Consequently, how should the population be able to assess its effectiveness? Thus, judgments made by the unemployed about its effectiveness are not very meaningful at this moment in time. For the unemployed, *ethics*, however, is ubiquitous in every regard of profiling. As decision making is increasingly delegated to machines and algorithms, new forms of discrimination by using data such as race, health status, etc. to inform automated decisions arise and moves ethics to the public focus (Pelzel, 2017). Through the simple categorization of people as a source of social stigma, some seem to appear better or worse than others (Niklas, et al., 2015). An inherent problem occurs when categorization of individuals as a method of social management takes place. Subsequently, there is a permanent conflict between those executing the profiling and those affected by it.

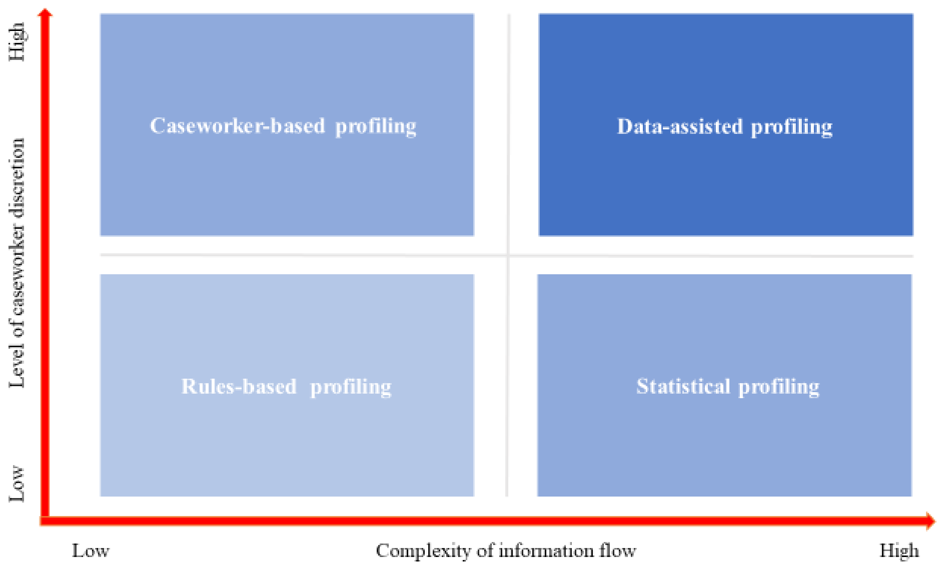
**Conclusion**

To conclude, the teacher could ask the students *“what would be the best way to design the organization?”* Thereby, he would encourage students to connect the dots throughout the different disciplines mentioned in the case and ensure that the quintessence is captured. A decision in favor of or against data-driven unemployment measures such as profiling represents a complex question when considering several stakeholders and their diverging interests. Hence, there is no such thing as one correct solution. A way to approach this question is to determine who holds the decision-making power and how the different interests are prioritized. Like this, a tendency towards or against data-science can be identified easily.

**Appendix 1:** EU Unemployment rate. % of unemployed in 2018, in *OECD Data*




**Appendix 2:** Approaches to risk profiling, in *Tackling Long-Term Unemployment through Risk Profiling and Outreach, 2018*



**Appendix 3: Data Science Project Scoping Worksheet**

1. *Project Name*
2. *Organization Name*
3. *Project Description*
4. *Who are the agencies/ departments that will need to be involved?*
5. *Who are the individuals in these organizations that are stakeholders? What are their roles?*
6. *Goals (in order of priority)*

*What are you maximizing or minimizing?*

*Are there any constraints (budget, resources, etc.)?*

| | | |
|---|---|---|
| *Goal 1:* | *Goal 2:* | *Goal 3:* |
| *Constraint:* | *Constraint:* | *Constraint:* |

7. *Actions*

*What is the action?*

*Who is taking the action?*

*What/ Who is being taken on?*

*How often?*

| | | |
|---|---|---|
| *Action 1* | *Action 2* | *Action 3* |
| *Questions* | *Questions* | *Questions* |
| *A.* | *A.* | *A.* |
| *B.* | *B.* | *B.* |
| *C.* | *C.* | *C.* |
| *D.* | *D.* | *D.* |

8. *Data*

*A. What Data do you have internally?*

| | | |
|---|---|---|
| *Data Source* | *Data Source* | *Data Source* |
| *What does it contain?* | *What does it contain?* | *What does it contain?* |

| | | |
|---|---|---|
| *What level of granularity?* | *What level of granularity?* | *What level of granularity?* |
| *How frequently is it collected/ updated?* | *How frequently is it collected/ updated?* | *How frequently is it collected/ updated?* |
| *Does it have unique identifiers that can be linked to other data sources?* | *Does it have unique identifiers that can be linked to other data sources?* | *Does it have unique identifiers that can be linked to other data sources?* |
| *Other* | *Other* | *Other* |

*B. What data can you get externally and/ or from public sources?*

| | | |
|---|---|---|
| *Data Source* | *Data Source* | *Data Source* |
| *What does it contain?* | *What does it contain?* | *What does it contain?* |
| *What level of granularity?* | *What level of granularity?* | *What level of granularity?* |
| *How frequently is it collected/ updated?* | *How frequently is it collected/ updated?* | *How frequently is it collected/ updated?* |
| *Does it have unique identifiers that can be linked to other data sources?* | *Does it have unique identifiers that can be linked to other data sources?* | *Does it have unique identifiers that can be linked to other data sources?* |
| *Other* | *Other* | *Other* |

*C. What data would you need in addition to the ones above?*

*Data Source:*

9. *Analysis*

*What analysis needs to be done?*

*How will you validate the analysis?*

| | | |
|---|---|---|
| *Analysis 1:* | *Analysis 2:* | *Analysis 3:* |
| *Analysis type:* | *Analysis type:* | *Analysis type:* |
| *Which action will this analysis inform?* | *Which action will this analysis inform?* | *Which action will this analysis inform?* |
| *How will you validate this analysis?* | *How will you validate this analysis?* | *How will you validate this analysis?* |

## Appendix 4: Board Plan and Order

Segmentation into *government* (top row), *IEFP* (middle row), and the *unemployed population* (bottom row)

| | | |
|---|---|---|
| Does the government correspond adequately to unemployment with profiling? | What are alternatives for the government to combat unemployment? | What is the role of the government? To which extent does its influence reach? |
| What are key challenges of managing profiling? | Why does IEFP use profiling? | What are reasons against the profiling model? |
| How does profiling match with the needs of the unemployed? | How does LTU affect the unemployed population? | What are perspectives on the increasing influence of data science? |

## Appendix 5: Potential Questions for Class Discussion – Class Flow for 90-Minute Session

*Block 1: The Government´s Point of View (25 min)*

| Key Questions | Sample Answers |
|---|---|
| 1. On which aspects do government bodies put their emphasis while fighting LTU?<br><br>2. Why did the government decide that profiling is the way to go?<br><br>3. Which (ethical) challenges appear in this regard? | - early intervention and proactive support of high-risk individuals, efficiency, timing<br><br>- legal requirements, best practices in other countries, advent of big data<br><br>- acceptance of machine-driven decision making, fair allocation of scarce resources |

*Block 2: IEFP´s Point of View (40 min)*

| Key Questions | Sample Answers |
|---|---|
| 1. Which are the most important attributes of the profiling model? On which features does IEFP focus?<br><br>2. Which features are missing or improvable in the current profiling model? | - precision & accuracy, categorization of individuals, early intervention, efficient allocation of resources, involvement of counselors<br><br>- transparency, easy-readable explanations, robustness against changing temporal conditions, dynamically generated risk scores, evaluation metric |

| | |
|---|---|
| 3. Which limitations is IEFP facing when operating its profiling model? | - dependence on government, organizational and technological maturity, conflict with counselors, confrontation with ethical biases |
| 4. How can IEFP make the model more impactful and respond to each of its stakeholders? | - continuous adjustment to latest developments in data science, implementation of evaluation metric<br><br>- being independent of government, extend collaborations *(data sharing)*, improve internal & external communication, prevent discrimination *(along the data value chain)* |

*Block 3: The Unemployed Population´s Point of View (25 min)*

| **Key Questions** | **Sample Answers** |
|---|---|
| 1. How is LTU affecting the unemployed? | - unexploited economic potential, poverty, negative psychological conditions (depression, somatization) |
| 2. What are the critical points regarding the profiling model? | - transparency, interpretability, effectiveness, ethics (omnipresent) |
| 3. How can discrimination be reduced to a minimum? | - equal distribution of resources, minimize categorization and increase transparency |

**Appendix 6: Assignment Questions**

By using different levels of abstraction, the teacher would have the flexibility to choose from a profound pool of questions, depending on the prevalent capabilities of students who get assigned to the case. Moreover, he could distinguish outstanding students from others.

| **Area** / **Level of difficulty** | **Rationale** | **Approach** | **Challenge** | **Evaluation** | **Project Scoping** |
|---|---|---|---|---|---|
| **Basic** (Descriptive) | What are drivers for IEFP to invest in data science? | What are different ways of applying profiling to help the unemployed? | Which challenges can you identify for IEFP to include data science? | How does the current evaluation of profiling look like? | Identify IEFP´s current project scoping framework. |

| | | | | | |
|---|---|---|---|---|---|
| **Advanced** | Is data science needed and suitable for combating unemployment? | What are criteria for choosing one approach over the other? | How would you prioritize those challenges? | Which additional KPIs would you define for the profiling model? | You have a task to develop a new data science project with the IEFP, using the project scoping experiences. Given the available information and the presented scoping framework, describe the project idea. |
| **Difficult** | Think about a different industry where data science has a significant impact? Are there similarities to the IEFP case? | Outline different scenarios in which each approach would be most suitable. Based on that develop fundamental pros and cons. | How would you overcome those challenges? | What changes need to be implemented at IEFP to make data-driven decision making sustainable? | |

**Teaching Note References**


Barnes, S., Wright, S., Irving, P., Deganis, I. (2015). *Identification of Latest Trends and Developments in Methods to Profile Jobseekers in European Public Employment Services: Final Report.* Published by: European Commission

Das, S. R. (2016). *Big Data´s Big Muscle.* Published by: International Monetary Fund

Data Science for Social Good (DSSG) (2018). *Predicting the risk of long-term unemployment in Continental Portugal: Final Report.* Published by: (n.d.)

Data Science for Social Good (DSSG) (n.d.). *Scoping Data Science (for Social Good) Projects.* Retrieved from https://dssg.uchicago.edu/2016/10/27/scoping-data-science-for-social-good-projects/

Eichhorst, W., Neder, F., Tobsch, V., Wozny, F. (2015). *A European Perspective on Long-Term Unemployment.* Published by: IZA Discussion Paper Series

Farber, H. S., Rothstein, J., Valletta, R. G. (2015). *The Effect of Extended Unemployment Insurance Benefits: Evidence from the 2012-2013 Phase-Out.* Published by: American Economic Review

Guerrero, O. A., Lopez, E. (2016). *Understanding Unemployment in the Era of Big Data: Policy Informed by Data-Driven Theory.* Published by: Wiley Periodicals

International Labour Organization (ILO) (2014). *Portugal: Tackling the Jobs Crisis in Portugal.* Published by: Geneva

Kotarba, M. (2018). *Foundations of Management: Digital Transformation of Business Models.* Published by: Sciendo

Kroft, K., Lange, F., Notowidigo, M. J., Katz, L. F. (2014). *Long-Term Unemployment and the Great Recession: The Role of Composition, Duration Dependence, and Non-Participation.* Published by: (n.d.)

Loxha, A., Morgandi, M. (2014). *Profiling the Unemployed. A Review of OECD Experiences and Implications for Emerging Economies.* Published by: Social Protection & Labour Discussion Papers

Nagy, H. (2016). *Mastering Digital Transformation: Towards a Smarter Society, Economy, City and Nation.* Published by: Bingley – Emerald Group Publishing Limited

Niklas, J., Sztandar-Sztanderska, K., Szymielewicz, K. (2015). *Profiling the Unemployed in Poland: Social and Political Implications of Algorithmic Decision Making.* Published by: Panoptykon Foundation

Nye, J.S., Zelikow, P., King, D. C. (1997). *Why People Don´t Trust Government.* Published by: Harvard University Press

Pelzel, F. (2017). *Big Data will be biased, if we let it.* Published by: Towards Data Science

Rosholm, M., Svarer, M., Hammer, B. (2004). *A Danish Profiling System.* Published by: Aarhus University Economics Department

Rudolph, H., Konle-Seidl, R. (2005). *Profiling for Better Services.* Published on behalf of the EU Commission

Scoppetta, A., Buckenleib, A. (2018). *Tackling Long-Term Unemployment through Risk Profiling and Outreach.* Published on behalf of the European Commission, Directorate General for Employment, Social Affairs and Inclusion by the ESF Transnational Platform