ELA
EUROPEAN LABOUR AUTHORITY

European Platform
tackling undeclared work

Subgroup on alert mechanisms to identify undeclared work at an early stage

Output paper

December 2022

Contents

Executive summary

The aim of the Platform *subgroup on alert mechanisms to identify undeclared work at an early stage* is to explore how relevant enforcement authorities can create alert mechanisms to identify undeclared work at an early stage. The meeting of the subgroup was held on 12 October 2022 in Bratislava (Slovakia) and online. This hybrid event brought together 35 members and observers of the Platform tackling undeclared work (the Platform), representing enforcement authorities from 17 countries and social partners. Representatives of the European Labour Authority (ELA) and the European Commission were also present. The members of the subgroup aimed to define the various types of alert mechanisms and databases used, identify opportunities and challenges involved in using such databases, exchange examples of good practices, and explore the range of more sophisticated forecasting methods.

Key messages

▶ Alert mechanisms can be described as procedures to identify indicators of growth, decline or development of new forms of undeclared work.

▶ The key elements for creating alert mechanisms include complex IT infrastructures with broad data sources, advanced methods of data mining, ability to identify changes in undeclared work patterns and expertise in predictive modelling.

▶ Some of the main data sources that can be used for the purposes of alert mechanisms include, but are not limited to, inspections records and outcomes, enforcement authorities' databases, complaints logs and population-based surveys on undeclared work.

▶ Currently, alert mechanisms are typically used in limited sectors. Their expansion to other sectors is hindered by legislative restrictions related to data sharing, and by limited cross-institutional cooperation.

▶ Enforcement authorities have developed various types of alert mechanisms. One promising example applies quantitative and qualitative risk assessment tools to identify companies for inspection based on a 'red-flags' system.

▶ Using a data-driven approach which employs artificial intelligence and text analysis in risk assessment provides opportunities to create alert mechanisms based on the efficient detection of patterns and signals of undeclared work.

▶ Alert mechanisms based on aggregate data, developed for instance by specialised analytical units, not only provide the possibility of integrating various databases but also help uncover various types of fraud, not only limited to undeclared work. This requires a joined-up governmental approach.

▶ Forecasting undeclared work requires advanced methods of data mining and predictive modelling. This can be achieved by combining different datasets and developing/using various databases with high integration and accessibility.

▶ Suggested next steps could include expanding the alert mechanisms already in practice to other sectors, increasing data sources, exploiting new technologies, ensuring adequate staffing and resources, and improving collaboration across institutions.

1.0   Introduction

The subgroup was established to bring together the knowledge created by the Platforms' previous studies on warning signals and risk assessment of future trends in undeclared work[1].The objectives of the subgroup are to:

▶ Define what alert mechanisms should identify, e.g. sectors, occupations, geographical areas, population groups where undeclared work is growing, specific types of undeclared work that are growing, and new emergent forms of undeclared work.

▶ Identify the databases enforcement authorities can use, such as (i) records of inspections, (ii) databases such as employment registers, tax return data and social insurance datasets, (iii) surveys, and (iv) notifications and complaints lodged by workers and other individuals/organisations, as well as exploring what type of data is gathered/used in these databases.

▶ Identify the type of aggregate-level analyses that could and have been conducted to identify trends.

▶ Determine the opportunities and challenges involved in using such databases to conduct aggregate-level analysis and gaps in available data.

▶ Discover examples of good practices in Member State enforcement authorities related to the application of alert mechanisms.

▶ Explore the range of more sophisticated forecasting methods (e.g. 'future basing', scenario-building, simulations, experimental laboratory studies of different scenarios and behavioural analyses, the use of artificial intelligence and machine learning) and the opportunities and challenges in using them in enforcement authorities.

▶ In doing so, the intention is to encourage enforcement authorities to share learning on how future-oriented risk analysis is, and could be, undertaken so that capacity building can occur.

The meeting of the subgroup held in October 2022 in Bratislava followed the above objectives. Additionally, discussions took place on the next steps in designing alert mechanisms. The meeting consisted of an introductory session with an intense discussion and presentation of the input paper, followed by several presentations on alert mechanism practices, available data and innovative ways of data mining, and risk assessment of undeclared work. Examples of sophisticated prediction methods and integrated data systems were introduced. At the end of the meeting, the participants formulated their 'wish list' to move forward in designing effective alert mechanisms.

This output paper documents the shared knowledge and results of the discussion of the meeting, and draws on the background information included in the discussion paper prepared prior to the subgroup meeting.[2]

---

[1] Williams, C., Horodnic, I.A. (2020) and Stefanov, R. (2021).
[2] Holubová, B. (2022). Platform subgroup on alert mechanisms to identify undeclared work at an early stage. Input paper for the meeting on 12 October 2022 (Bratislava) (unpublished).

2.0 Current knowledge of alert mechanisms

Alert mechanisms are procedures to identify the determinants/factors of growth, decline or emergence of new forms of undeclared work. Alert mechanisms enable enforcement authorities to be proactive by intervening early on at both the operational and policy levels to prevent the growth of undeclared work. Alert mechanisms allow enforcement authorities to apply targeted, timely actions and thus save human and financial resources.

Box 1. The key elements of alert mechanisms
▶ Complex IT infrastructure of integrated and accessible data related to undeclared work;   ▶ Advanced methods of data mining, risk analysis and forecasting;   ▶ Ability to identify/set up the determinants/factors contributing to the development (changes) of undeclared work;   ▶ Expertise in data mining, risk assessment and predictive modelling.

Sources: Williams and Horodnic (2020), Stefanov (2021) and the shared knowledge at the subgroup meeting

Alert mechanisms can be applied at the micro-level and target specific sectors, occupations, companies, regions or types of undeclared work. They can also assess the risks of undeclared work at a macro-aggregated level and predict broader trends in undeclared work.

To create alert mechanisms, four types of data can be used to assess the trends and detect early signals:

▶ Records of inspections and their outcomes (inspection registers or case management databases);

▶ Enforcement authority databases (employment registers, tax return data, and social insurance datasets);

▶ Complaints and notifications by workers (e.g. collected by inspectorates or social partners);

▶ Surveys of undeclared work (e.g. Eurobarometer surveys on undeclared work, European Working Conditions Survey).

Some countries have established specialised units for centralised data collection and data mining. Such dedicated departments are found in Norway (The Directorate of Tax), Spain (The Anti-Fraud Tool Unit) and Finland (The Grey Economy Information Unit). Based on the previous studies carried out by the Platform, several indicators – red flag signals - of undeclared work have already been identified, such as past infringements, below average salary, a high number of short-time contracts, etc. (Table 1). The red-flag signals can be considered when developing alert mechanisms to detect undeclared work at an early stage.

Table 1. Red flag signal of undeclared work by type of illegal employment

Unregistered employment	Under-declared employment	Bogus self-employment
Turnover and number of employees mismatch; Past infringements, fines, penalties; Newly established companies; Companies in high-risk sectors (e.g. construction); Firms/sectors cited in the received complaints/signals; and Firms/sectors identified as risky by inspectors and other authorities 'on the ground'.	Below average salary for the sector/profession; High proportion of workers paid minimum salaries; 25-50 % are part-time employees; Part-time contracts are less than 4 hours/week; High number of short-time contracts; Past infringements, fines, penalties; Firms/sectors cited in the received complaints/signals; and Firms/sectors identified as risky by inspectors and other authorities 'on the ground'.	Sudden changes from employees to self-employed (continue working for the same entity paying their wages); Past infringements, fines, penalties; Newly established companies; Large number/quickly changing subcontractors; Large numbers of 'go-betweens' without license acting as temporary work agencies; High turnover of staff (more than 10 % per month) though there is a low overall number of workers; Companies in high-risk sectors (e.g. construction); Firms/sectors cited in the received complaints/signals; Firms/sectors identified as risky by inspectors and other authorities 'on the ground'.

Source: Williams, C., Horodnic, I.A (2020). Horizon scanning: early warning signals of future trends in undeclared work European Platform tackling undeclared work.

ELA's coordination role is irreplaceable in creating effective alert mechanisms. In this regard, the **Labour Mobility Analyses and Risk Assessment (LMARA)** team is relevant. Its objectives include carrying out analyses and risk assessments on issues of cross-border labour mobility within ELA's remit and social security coordination across the Union. Supporting cooperation between Member States (analyses, risk assessments, data sharing) and exchanging experience and promoting capacity building are additional tasks of LMARA.

The main domains of analytical and risk assessment capacity of LMARA are:
- ▶ Operational risk assessments providing analytical support to concerted and joint inspections (CJIs),[3] pro-active search of companies and analysing inspection outcomes, processing data, suggesting follow-ups, etc.

---

[3] More information about the CJIs available at: https://www.ela.europa.eu/en/concerted-and-joint-inspections.

- ▶ Strategic risk assessments focusing on sector-specific challenges, recurring problems at national and EU levels, focused in-depth analyses and studies to investigate specific issues.

- ▶ Analyses of policies and regulations in the Member States and fostering competent authorities' analytical capacities.

- ▶ Outreach and capacity building encompassing the capacity building workshop, the network of experts/analysts and peer reviews on application of Union law within the ELA's competence.

The following section presents diverse alert mechanisms developed and used by enforcement authorities, their main elements, success factors and related challenges. The text is based on the presentations at the meeting and the resultant discussion.

3.0   Shared learning on alert mechanisms

A flash survey among the participants revealed that most of the participating authorities use some kind of alert mechanism based mostly on data from inspection registers and case management databases. Nevertheless, the participants reported encountering several challenges.

Alert mechanisms in place only in specific sectors

The State Labour Inspectorate of the Republic of Lithuania uses alert mechanisms based on the risk assessment tools to identify risk companies in occupational health and safety matters, undeclared and under-declared work. However, so far, the tool has only proven effective in detecting undeclared work in the HORECA sector. Whereas, in the construction sector, which is particularly affected by undeclared work, the risk assessment tool is difficult to use due to challenges in identifying the address of construction sites, as in some instances the authorities have access to the address where the business is registered, but not necessarily the address where workers perform the construction work.

Legal challenges to data sharing

The Swedish Work Environment Authority has implemented alert mechanisms based on identifying the risk companies using artificial intelligence (see the section 3.3.). However, problems remain concerning data sharing, including between the responsible authorities, due to strict legislation in place, and the fact that the system is requires manual input and is therefore time-consuming. In addition, the Central Customs Authority in Germany is facing legal challenges in integrating, building and combining the needed datasets for risk assessment purposes and is therefore working on a new legal basis as well as a legally compliant technical solution.

Limited cooperation with the local labour inspectors

The Hellenic Labour Inspectorate experiences the involvement of local labour inspectors who need to undergo special training on how to use the necessary tools. Data from the ERGANI database could be used to predict undeclared work trends at an aggregate macro-level. However, there are challenges related to limited human resources. Similarly, the National Labour Inspection in Poland stated that the local labour inspectors need to understand why it is essential for them to insert comprehensive data into the databases, and how some of the colleagues could use that data. However, there is a lack of trust between companies, individuals, and the state that hinders the process.

Lack of analytical system despite developed cooperation

The Ministry of Social Affairs and Health in Finland reported a lack of analytical systems and analysts in the occupational safety and health (OSH) administration, despite multilateral cooperation and data sharing among authorities. The representative of the Ministry of Labour and Social Affairs in Czechia indicated they have bilateral agreements in place between the labour inspectorate and the financial directorate on the transfer of information related to illegal work. They share knowledge through inter-ministerial bodies for tackling undeclared work, which meets twice a year.

Promising new data sources to be exploit for alert mechanisms

The Department of Labour Inspection in Cyprus uses information from a hotline on undeclared work and unregistered employment to identify trends; this whistle-blower system is anonymous. The Financial Administration of Republic of Slovenia uses the fiscal cash register in all sectors, whereby one must submit invoices. The system is connected to the tax authority which collects data on the employer who issued the invoice in the system; with that data, they can check if the person is legally employed.[4]

The Hellenic Labour Inspectorate, Swedish Work Environment Authority and National Labour Inspection in Poland shared their experiences with designing alert mechanisms in detail to inspire other countries but also to discuss the challenges and advantages of the tools.

3.1   Red-flags system in Greece

During the meeting, the Hellenic Labour Inspectorate described the design and subsequent experience they had acquired of an **alert mechanism based on applying a risk assessment tool to identify companies for inspection based on a 'red-flags' system**. The aims of the risk assessment tool are multiple. Firstly, the tool identifies groups of companies with specific characteristics linked with a high risk of delinquency[5] and consequently targets these companies with focused awareness-raising campaigns, sending notification letters, onsite inspections, monitoring of employment declarations or other initiatives. Secondly, the tool increases the effectiveness of applied measures and, thirdly, seeks to successfully allocate the human and financial resources of the labour inspectorates.

The Risk Assessment tool is a subsystem of the Integrated Information System of the Labour Inspectorate and uses data from the ERGANI Information System[6] and the Inspections subsystem (IIS). It is configured at the Labour Inspectorate's central offices by risk analysis experts, based on labour inspection expertise on high-risk/red-flag indicators.

For each **red-flag indicator**, one risk analysis 'rule' is created. A risk analysis 'scenario' is comprised of multiple weighted rules combined. Each risk analysis scenario may be generated to target a specific area, business sector, type of infringement, etc. Once a risk analysis scenario has been configured and tested, it may be used for

---

[4] The fiscal cash register is used by more countries, and there is a potential to use the record for creating alert mechanisms.
[5] The term delinquency is used here as a general term. Despite that the risk might relate to administrative infractions, there could be a risk of delinquency related to criminal infringements. Undeclared work could be linked to labour exploitation, tax fraud non-compliance with OSH regulations, which in most cases are considered criminal infringements.
[6] The ERGANI in Greece is a centralised employment data system that aims for more data sharing and access for different authorities. The system has established interoperability connections with tax and social security authorities' systems to provide employers with several online services for declaring all their employment details before these are put in place. However, no steps have been taken in the field of data matching or data mining to use these authorities' data with a view to tackling undeclared work.

assessing the companies' level of risk. Using various selection criteria, a list of companies in a specific area and/or business sector is created, and their risk level is calculated. Followingly, a list of high-risk companies is submitted to local departments by automatically initiating an inspection case for each of them in the Inspections subsystem to conduct onsite inspections.

One of the relevant features of this alert mechanism is that it may be **combined with a quantitative and qualitative assessment**. After generating the list of high-risk companies by applying the indicators, the local inspectors may be asked to apply their own risk assessment based on their experiences with the companies in question before the onsite inspections take place. **Employing local inspectors' knowledge,** which the central level does not possess, is an additional important characteristic of the alert mechanisms.

When **piloting the tool**, undeclared and under-declared work in HORECA and personal services were targeted. The risk analysis rules and scenarios were created based on specific red-flag indicators[7] and ranked the businesses in the targeted sectors. Four different local departments were involved, and the labour inspectors were trained and fully informed of the project objectives and the details of implementation. Three types of notification letters were sent to the 400 riskiest companies in each sector and area ('gentle nudge' letter, 'strong nudge' letter, announced inspection letter). The lists with the 400 riskiest companies were delivered to the local departments to perform onsite inspections after selecting half of them using their own local criteria.

The evaluation of the targeted inspections confirms a positive impact. There was an increase in undeclared work detection compared to the previous year in three out of the four local departments. In the areas where all four pilot local departments were located, there was a significant increase in full-time contract declarations with a concurrent decrease in part-time contracts – -which can be translated as a transformation of part-time contracts into full-time– and an increase in declared overtime when compared with the rest of the country.

Success factors of the Greek red-flag assessment tool stem from efficient cooperation and communication between the risk analysis experts and the local labour inspectors with regard to identifying indicators per targeted delinquency type in a particular local area. Moreover, the local inspectors were involved in the process of planning. The proper configuration of the Risk Assessment tool, the continuous evaluation of its effectiveness, and the inclusion of targeted onsite visits in the annual and monthly inspection plans are key factors contributing to the tool's effectiveness.

Challenges include mainly the lack of experienced personnel at the Central Offices dedicated to configuring the scenarios and evaluating results concerning the well-functioning of the tool. So far, the tool has been used for micro-level risk assessment to identify local cases. Potentially, the same red flag indicators can also be used on a macro-level to identify where undeclared work is more likely to occur and what types are growing over time. Though the ERGANI system contains millions of records of employment data from employers, which can be put to further use, the lack of specialists to process the available data makes this level of analysis of emerging trends unfeasible. Another challenge concerns the inclusion of locality and seasonality factors in risk analysis scenarios to address the diversity in the labour market countrywide.

---

[7] The indicators generated for this case were a high number of recruits and layoffs, fines for undeclared or under-declared work within the last two years, a high number of part-time employees, a high number of changes in the employees' contracts from full-time to part-time and an increased number of changes in the employees' working hours.

3.2 Fighting the 'first day at work' syndrome in Poland

The National Labour Inspection of Poland (NLI) presented the **alert mechanisms which had been created after the inspection records frequently indicated the 'first day at work' excuse.** The 'first day at work syndrome' phenomenon occurs when employers - who illegally employ workers without an employment contract - would claim, backed by the worker, that it is the worker's first day during inspection visits. This problem originates in the former Labour Code, in force up until 2016, as it allowed for the employment contract to be written at the latest on the date the employee started working. However, in case of an inspection, both employer and employee would falsely state that the first day of work was the same date as that of the inspection.

The inspection records from 2015 indicate that 10 000 people said they started work on the day of the inspection, i.e. without an employment contract written down and without registering for social insurance. Moreover, 1 400 companies claimed that the employee in question started work on the inspection day. To address this discrepancy, the enforcement authorities selected businesses for inspections in which illegal employment was identified in previous years comprising micro and small enterprises in sectors such as trade, industrial processing, construction, HORECA and enterprises engaged in work based on civil law contracts. The selection was based on signals reaching the NLI, comprised of complaints, press reports, social media messages and information based on cooperation with other authorities.

The 2016 amendment of the Labour Code, initiated from the NLI, introduced the obligation to conclude the employment contract in writing before permitting the employee to perform their work duties. Based on the NLI's usage of the alert mechanisms, the number of irregularities has been decreasing. While in 2016 illegal employment was revealed during 9 300 inspections affecting 25 200 people by the non-compliance of with regulations, in 2017, illegal employment was noted during 5 600 inspections affecting 13 800 people, including 2 600 employees who did not have a written employment contract and 12 800 people who were not registered for social insurance.

Lessons learned:

▶ Establishing a prior notification system by the labour inspectorate could form the basis of an alert mechanism that can be built upon by adding other mechanisms.

▶ The limited remits of the enforcement authorities resulted in frustration for the inspectors, and at that time the only way to decrease irregularities was to change the respective legislation.

▶ The lack or insufficient definition of undeclared work hampers the design of precise alert mechanisms with set indicators.

▶ In case of the absence of a clear definition of undeclared work, only the infringement of a specific obligation or regulation, such as non-existent working contracts in writing, can provide proof of irregularities.

3.3 Artificial intelligence risk assessment in Sweden

Sweden shared the experience of **employing artificial intelligence (AI) and text analysis in risk assessment.** This innovative means of risk assessment stems from the possibilities of new technologies to process a vast amount of data by artificial intelligence and machine learning. A lot of information is still in text form and uncategorised or is not even digitalised. The new technologies created conditions to find significant patterns by analysing the data differently with AI and text analysis.

This **data-driven approach** was initiated by the Swedish Work Environment Authority[8] in a project aiming to create a working environment that does not lead to sickness, injury or death. To achieve this vision, the analysis of previous accidents in the workplace serves to define the biggest risks within the work environment and identify patterns to be avoided in the future. The identified pattern of the high risk of accidents can then be perceived as signals at an early stage based on historical data, which opens up the opportunities to see these early signals in other working environments.

The AI risk assessment uses two **sources of data**. First, the incoming reports of the accidents are based on the obligation of the employers to report all work-related injuries, death and accidents that entail danger to life or health. An additional source of data is the incident reports, submitted by workers anonymously via an online form.

Firstly, they organised a **crash course** with key persons and stakeholders within the Authority to learn more about AI and how they can use it. They then dig deeper into their data and different AI methods. The crash course was the start of spreading knowledge within the organisation.

Secondly, they **explored the possibilities and challenges** of how the inspectors manually processed, sorted, prioritised, and engaged with the incoming reports of personal injuries, accidents or death related to the workplace.

Thirdly, they **developed prototypes** based on prioritisation and detect patterns. This chosen AI-solution helps prioritise and detect patterns in incoming reports on accidents. It also identifies keywords or phrases in a report, such as the number of victims, number of sick leaves, categories etc. The prioritised list of the report is sent to inspectors for review and to form a basis for the decision. The AI system then analyses the reports and decides whether it is suitable for the risk assessment category or not. The AI system, based on the text analysis, sorts out incomplete reports, helps to act fast and on the right signals and saves time for the inspectors.

The project is currently in the **data readiness stage.** This means digging deep into the data, checking if it is ready and suitable for use, and how it should be prepared for further analysis.

Box 2. Advantages of the AI risk assessment

▶ Saving time for inspectors;

▶ Ensuring uniform and standardised processes in a faster way;

▶ Fewer people reading classified information;

▶ The final decisions are made by humans.

The prototype developed is the first proof of concept and will be used to prevent work-related accidents and other risk patterns to ensure occupational health and safety. Potentially, a **similar concept can be developed to detect the patterns and signals of undeclared work**. The question is which data source(s) can be used, accessed, and shared. Sweden has a similar initiative encompassing cooperation among eight authorities against crime connected to the workplace, including the Swedish police and the Public Employment Service. This allows for the

---

[8] The Swedish Work Environment Authority is a regulatory authority with the mandate from the government and the Riksdag to ensure that laws about the work environment and working hours are followed by public and private companies and organisations.

sharing of relevant data. The initiative is still in preparation mode with a view to scaling up. The idea is to start small and to use AI technology to see what it can do.

To ensure the correct (valid) set of information is inserted in the system, to avoid biases and increase trust in the outcome, proof-checks at multiple layers are used to eliminate the mistakes in the algorithms. There is a double proof at the beginning of the process, both manual-human assessment of the reports and AI-based data analysis of the same sources. This double proof is a sufficient test of the validity and precision of the AI system.

The use of AI is often associated with the idea of **high costs**. Not in this case. The cost of the AI machine learning system in this project is relatively affordable as only pre-trained models from the National Library have been used as opposed to creating new models from scratch. Currently, the only cost of the project comprises advanced computers capable of employing the AI models and the salaries of two digital experts.

Lesson learned from the case study:

▶ Start small and scale up – do not get caught in a phase of pre-studies;

▶ Solve problems when they occur – do not make up hypothetical problems;

▶ Competence – choose a strategy but start the competence improvement today;

▶ Start with a pre-trained AI model – it is an excellent platform to be built on and saves costs.

4.0 Alert mechanisms at the aggregated level

The aim of the session was to discuss the role of data, access to relevant datasets and types of data used in designing alert mechanisms. Spain presented a complex centralised data system that allows for the application of analytical tools to detect undeclared work. After that, the limits in accessing and sharing data were discussed. The outcomes and suggestions for solutions are summarised in the second half of this section.

4.1 Centralised tool to detect fraud in Spain

Spain introduced the **Anti-Fraud Tool Unit** and examples of the analysis tools used to predict and prevent fraud in employment and social security. The Anti-Fraud Tool (HLF) was established as a Unit of the Labour and Social Security Inspectorate (ITSS) in 2015 but was only implemented in 2018. The HLF, through the use of data analysis techniques, detects cases that signal indications of fraud and selects them for further planned actions and assessment. Probabilities are calculated based on business models and are backed by data mining. The staff of the HLF comprises inspectors, sub-inspectors and IT officers as well as external staff.

The core of the HLF is the SAS platform administrated by the unit staff. The SAS platform is a software that seeks to extract valuable knowledge from unstructured data. It also has the ability to process information automatically through machine learning. The HLF uses the SAS branch to collect, manage and process data and distribute suspected fraudulent cases. The HLF is interconnected to the regional inspectorates via INTEGRA, software for data sharing. The cycle of an integrated process between the central level and regional inspectorates consists of

issuing the specific fraud model and selecting subjects to be inspected at the central level. The complex files on the subjects are then loaded into INTEGRA. The regional offices, based on the files, assign officers with service orders to inspect the cases. The results of the inspections are uploaded to the system, and the HLF evaluates the effectiveness of the model (Figure 1).

Figure 1. The cycle of integrated processes of the Anti-Fraud Tool Unit
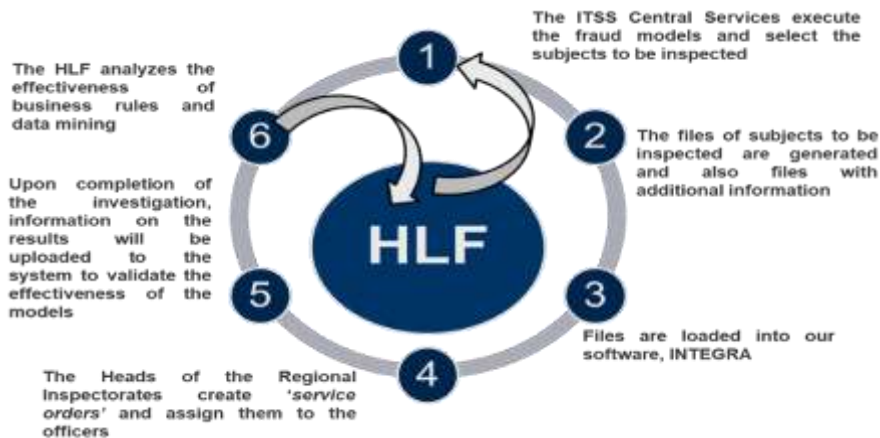
Anti-Fraud Tool Unit, (HLF) Spain. Presentation at the subgroup meeting for Alert mechanisms to identify undeclared work at an early stage on 12 October 2022 (Bratislava).

Thus, the HLF serves as an instrument to plan actions at the national and regional offices. To this end, a relational data mart has been created that allows for the designing and activation of the following actions:

- ▶ Business Rules integrating National Anti-Fraud Office, Labour Relations and OSH;

- ▶ Predictive Models integrating National Anti-Fraud Office and Labour Relations and OSH;

- ▶ Activation of Extensive Actions, called Shock Plans.

The **business rules** are processes that parameterise data according to requirements indicative of fraud and allow for the selection of potentially fraudulent cases. New business rules can be developed based on the accumulated knowledge by the ITSS, new instances of fraud detected in the actions of the ITSS or the study of databases.

The HLF business rules on social security and undeclared work can relate to diverse types of fraud:

- ▶ Control of the registration of workers: for this area, different rules have been developed with regard to the registration of workers and with the goal of tackling erroneous or fraudulent frameworks;

- ▶ Social Security contributions: several rules have been developed that analyse possible frauds generated in the area of Social Security;

- ▶ Debtor Companies: analyse the debtor companies to develop actions in this area;

- ▶ Fictitious Companies: these rules look for possible fraud in terms of access to benefits through fictitious registrations;

- ▶ Short-term financial support schemes control: in the last year, several rules have been developed to address possible fraud that may occur in companies benefiting from short-term financial support schemes;

- ▶ Social Security benefits: use of indications of possible fraud in terms of access to benefits.

Business rules related to labour relationships, equality and OSH can detect the following:

- ▶ Fraud in labour contracts: several rules have been developed that analyse possible fraud in contracts; most of them focus on the excess of the legal duration limit of temporary contracts;

- ▶ Wage discrimination (e.g. based on gender);

Occupational health and safety – accidents at work. An example of business rule is the model targeting bogus self-employed statuses, i.e. identifying companies that could have formalised labour relations with an employee without the corresponding registration, but the employee is registered as self-employed in the social security system, with the consequent saving of contributions falling on the self-employed person. This rule is based on external files provided by the Tax Authority related to people who present a high amount as self-employed.

Additionally, the HLF can apply **Analytical Models.** Based on the study of cases verified as fraudulent and using mathematical representations, analytical models aim to determine the existence of an infringing pattern which serves to detect companies that present similar fraud situations. Some analytical models have been developed which take into account the socio-labour characteristics of the territories so that they are more effective in parameterisation and fraud detection. An example of an analytical model might be the identification of letter box companies. The methodology involves the construction of variables that may be relevant to differentiating fictitious companies from those that are real, such as analysing debts with Social Security; the number of workers employed in the last four years; the number of days that workers have worked on average in the last four years; different sectors of previous employment, etc. Each variable has a specific 'weight' based on mathematical models.

The **shock plan** saw extensive action taken to detect fraud regarding temporary contracts. The aims of this action were to regularise fraudulent temporary contracts, provide stability to workers, deter non-compliant companies from further fraud and obtain better results and human resources savings. The shock plan has been shown to have a huge impact in a short period. The action is executed in two phases. Firstly, the HLF selects the companies that have allegedly infringed the law in the area of temporary contracts. Clear notification letters are then sent, which include a deadline to rectify the infraction. In the second phase, the HLF verifies in the database if the correction was made. If not, inspectors' investigations begin with regard to those companies that did not regularise the contracts following the notification letter.

4.2   Discussing data access and exchange

The alert mechanism and risk assessment depend on the quality of the complex data and their analysis. The enforcement authorities have limited access to necessary administrative registers, such as tax or immigration databases, when it comes to using them to tackle undeclared work.

The 2022 Platform survey on progress towards a holistic approach reveals that 8 % of authorities responding have limited access to data from other authorities. Meanwhile, 88 % have electronic access to some data from other enforcement authorities, but just 4 % have full electronic access to all relevant data from other authorities' databases. Sharing data, therefore, could be significantly improved.

Additionally, results from this survey suggest that none of the responding authorities has a central unit that holds all the data and undertakes the analysis for all authorities. Similarly, none of the respondents stated that they could directly analyse all relevant databases from other authorities. Only 21 % of respondents can directly analyse some

relevant databases from other authorities. 50 % receive data from other authorities that can be imported into their own databases, and 29 % can only directly analyse their own databases.

Challenges of data availability and data sharing:[9]

- ▶ Lack of political will/culture of exchange;

- ▶ Lack of understanding of data protection regulations;

- ▶ Missing operational processes;

- ▶ Exchange with tax authorities seems to be more limited, although relevant;

- ▶ Some countries lack a detailed legal framework and rely on informal cooperation;

- ▶ Different systems and targets among authorities;

- ▶ New data sources call for new cooperation solutions (e.g. social media and online data scraping);

- ▶ National vs local level coordination;

- ▶ The paradox of too much information and the difficulty in using it.

The meeting participants revealed challenges with data access and exchange and shared efficient solutions to overcome the obstacles:

Overcoming GDPR challenges

- ▶ In Lithuania, GDPR rules present a challenge, hampering the free movement of information between institutions. Conversely, data and information exchange is obligatory within Lithuanian institutions.

- ▶ It has been noted that some Member States have managed to overcome GDPR-based obstacles. The common framework of data protection in the EU allows the sharing of practices in place between Member States.

Combining datasets

- ▶ In Lithuania, many databases are available online, such as tax and social security databases. The problem concerns how to merge them into one database. Significant human and financial resources are required to solve this problem.

- ▶ Several participants suggested that not all enforcement authorities actually need to have access to all databases for data mining and designing risk assessment tools. A solution would be to have one analytical unit, a centre which would provide a risk assessment similar to that in Spain or Belgium (see the next section). This would solve the problem of the lack of IT experts and data scientists.

- ▶ The Central Customs Authority in Germany intends to data mine relevant datasets for risk assessment purposes. Portugal faces challenges with the confidentiality of several national databases, which leads to lack of informational exchange. Furthermore, exchange of information is not part of structured databases.

---

[9] Stefanov, R. (2021).

- ▶ The Central Customs Authority in Germany started data mining tax datasets for relevant data related to risk assessment. Portugal faces challenges with the confidentiality of several national databases, which leads to lack of informational exchange. Furthermore, exchange of information is not part of structured databases.

- ▶ Belgium proposed a country-specific solution to data access and sharing. They implemented institutions-based data exchange, the so-called Crossroads Bank. One must create 'authorisation demands', within which one must be clear about the repurposing of the used data. Based on the authorisation demand, one can get direct access to the data one needs. The Crossroads Bank is not about storing the data in a central system but ensuring a smooth data flow. This allows for decision to be made on data exchange without the need for political will.

- ▶ In Greece, the law of the Ministry of Digital Governance from 2019 regulates that all data must be transferred via the Inter-operability Centre. The authority governs the data exchange between public institutions – similar to the system in Belgium. On a publicly available website, one can check all the available data and databases.

- ▶ Poland's labour inspectorate has currently limited access to companies' turnover data from the tax authorities. A change of the legal framework granting the labour inspectorate permission to access this type of data, as well as close cooperation and trust building between the institutions would help address this challenge.

- ▶ Finland needs more political will to grant authorities access to the databases and registers. Only after this step is taken could a complex analytical system for alert mechanisms potentially be set out.

- ▶ The Authority for Working Conditions (ACT) in Portugal also faces difficulties in collecting information. However, there is some exchange with the social security services. They suggest asking for, or sharing, rather small files at the beginning.

Additional challenges related to data access and IT infrastructure:[10]

- ▶ Lack of access to each other's databases across authorities;

- ▶ Lack of knowledge, hardware and software for remote access to databases;

- ▶ Lack of budget, training, and human resources;

- ▶ Different inspection authorities rely on different unique identifiers, which hampers data matching and mining;

- ▶ The cost for contractors to implement interoperability is high;

- ▶ The amount of work needed to be invested in coordinating with other authorities on what, and how to, exchange;

- ▶ Need for more capacity to address ongoing technical challenges;

- ▶ Centralised systems do not automatically lead to more data-matching capacities.

---

[10] Stefanov, R. (2021).

5.0 Forecasting undeclared work

This section explores the range of sophisticated forecasting methods to detect undeclared work. Belgium shared some possibilities for detecting undeclared work through data mining and set out the possibilities provided by various types of centralised data collection. The aim of this meeting session was to think 'outside the box' when it came to data (e.g., future-basing, scenario-building, simulations, experimental laboratory studies of different scenarios and behavioural analyses, the use of artificial intelligence and machine learning) and the opportunities and challenges in using them in enforcement authorities

5.1 Predicting undeclared work by data mining in Belgium

Belgium contributed to the topic of alert mechanisms by sharing the **possibilities to detect undeclared work with data mining.** Belgium is a country with a well-developed, centralised system of data collection, data mining and data sharing. The centralised system is administrated by the National Social Security Office (NSSO) which has broad competencies. The NSSO collects social contributions based on the contributions on contract and labour agreements; funds the social security system; gathers, governs and distributes wage and career data; shares the data through the crossroads approach, fights social fraud, social dumping & human trafficking and provides policy support and analyses data.

At the same time, the NSSO is a driver for e-governance, striving for digitalisation, enabling data mining and simplifying processes. To do this, the NSSO applies several online applications for data collection, thus allowing for the planning of data mining, predictive modelling and forecasting future trends. The NSSO administrated the following databases:

- ► The social security card (SIS) to identify workers based on an e-ID card, including the possibility to identify workers by residence, has been in place since the 1990s;

- ► The DMFA collects multifunctional declarations from employers which include all details of workers' performance and compensation. This allows for the calculation of social contributions. The NSSO passes this data on to the Central Social Security Database. The database also shows the increase of employment in concrete economic sectors using the data on sectors;

- ► Checkinatwork - this monitors the presence of workers in the workplace (especially important in the construction sector at the construction sites), in addition to when they leave, also allowing for the planning of onsite inspections;

- ► DIMONA is an online service for employers to report the entry and exit of employees. The employer-employee relationship data is unique and critical to employee identification. The NSSO shares the information with other social security institutions to protect social rights;

- ► Limosa gathers information about posted workers and has administrated A1 Forms since 2007. A foreign employer who posts an employee temporarily or partially to Belgium must declare this activity in advance via the Limosa online service if the employee is not subject to Belgian social security. This is also the case for self-employed people who are based abroad but carry out an assignment in Belgium. Data is collected in the Limosa cadastre and shared with interested entities and inspection services;

- ► Studentatwork collects data on the place where the student work is done and also indicates the growth of sectors; in operation since 2012.[11]

Despite the centralised and high level of digitisation of data governance in Belgium, detecting undeclared work is not simple. In general, it is possible, though can be more complicated than expected as advanced data mining is needed. Moreover, on the inspection side, it might be even more difficult as typically one needs to consider many different variables and consider that undeclared work occurs on different levels and at different time scales.

A potential indicator of undeclared work occurs when companies with a large turnover that have fewer workers than competitors are detected. Hypothetically, a particular sector can be considered, with particular regard to the region and timeframe in question, and there is also a correlation between the annual turnover of companies, the workforce needed to carry out the work and the actual number of workers (FTE) etc.

The preconditions - enablers to detect undeclared work through data mining (not exhaustive):

- ► Combining different data; the more sources, the more precise the modelling;

- ► High level of e-government with regard to accessibility of data;

- ► Extensive business knowledge;

- ► Complex data integration & application integration – careful feature modelling (i.e. indicator, predictor variable);

- ► Complex investigation procedures with real-time data, or almost real-time data (this allows for extra possibilities concerning investigations);

- ► Building of predictors based on data from many different applications across institutions;

- ► Application of multivariate techniques, data mining, and predictive modelling.

One of the enablers of predictive modelling is **feedback registration.** Feedback registration means recording all the information – the data relevant to the case being investigated - in a structured way**.** The feedback registration can enable data mining based on supervised learning, where the key input is the confirmed case of fraud.

The feedback registration might be **short-term**, i.e. immediately after the inspection at the workplace. Possible results may include the detection of undeclared work, the violation of work conditions, the absence of legal documents, suspicion of fraud of a specific type, etc. The results of the **mid-term** feedback registration, i.e. after the conclusion of the inspection 'case', might be the confirmation of fraud type X or Y, etc., the issuance of declaration updates or the need for extra contributions, etc. Finally, the **long-term** feedback registration could include combining cases over time or across networks that take into account the verdicts of legal cases, the percentage of due contributions actually collected, etc.

To obtain quality feedback registration that can feed into predictive modelling, it should be as rigorous as possible, i.e. be followed up on, contain due dates and be supported by advanced IT applications. Moreover, instances of fraud must be categorised into clear types and time-stamped to develop predictive models. This way the dynamic can be better explored and new phenomena, such as new types of fraud or system abuse, can be detected.

---

[11] More information about the NSSO databases, see: https://www.lss.be/e-government

The data mining also proved to be beneficial regarding Return on Investment (ROI) and Cost Benefit Analysis (CBA). For example, the field action day tests in 2019 resulted in a 75% hit rate at the individual level. The action on detecting social dumping required approximately 150 field action days on work sites, and more than 500 companies were inspected yearly. After using appropriate data mining, detection accuracy went up from 35% - 50% to 70% - 85% at the work site level. In the action on failed carousels, based on improved mapping of the network and subsequent spatial visualisation of the most effective route, the work of inspectors was reduced from several days to several hours. Also, the companies top-ranked by predictive models have proven to bring twice as many positive results to detect regulations infringement and frauds, thus saving the cost and increase of the return on investment.

5.2   Additional forecasting methods

When analysing data, risk assessment processes typically use a linear extrapolation of the past to provide early warning signals to predict future events. For example, if unregistered employment is growing in the agricultural sector, it could be considered likely to do so in the future. However, the future is rarely a linear extrapolation of the past. Other and more sophisticated, forecasting methods and techniques exist there, that are heavily based on qualitative methods to predict undeclared work. Examples of the kinds of forecasting methods that could be used by enforcement authorities are provided below.[12]

Future basing – participants choose some point in time in the future (e.g. 15 years from now), develop a detailed picture of what this future world looks like and then identify the steps they need to take to achieve that world. They thereby reconstruct the detailed operational plan they created to move from the situation in 2022/23 to the current ideal world they now inhabit in 2040. The outcomes are then used to inform present-day strategy.

An example of future basing might be to think how the digitalisation of the workforce (home office, teleworking) will change the role and opportunities of the labour inspectorates in the near future. Imagine you are living in future: what does society and the economy in the EU look like in 10 years? In this future, work life and homelife are no longer separated, one increasingly works from home. What must be changed to prevent undeclared work in this context? How do tax authorities and labour inspectorates have to adjust to meet the requirements of the future?[13]

Scenario building – this can be described as creating a story about the future based on an analysis and understanding of current and historical trends and events. The development of sets of future narrative scenarios then enables labour authorities to engage in scenario planning.

Scenario planning, also called scenario thinking or scenario analysis, is a strategic planning method that some organisations use to make flexible long-term plans. It is a creative process, much like writing a novel, with a plot beginning with the current reality. The objective is to create a variety of plausible futures (scenario building) and then to analyse what the organisation would need to do to succeed under each of them. It is generally used to assess the risk associated with a key decision being considered.

Behavioural laboratory experiments and simulations – these have been extensively used by economic psychologists with regard to tackling undeclared work, especially tax compliance, and to evaluate how participants

---

[12] Williams, C., Horodnic, I.A (2020).
[13] Intervention of Collin Wiliams at the meeting.

will respond when specific new policy initiatives are introduced (e.g. higher penalties, a greater risk of detection, being told by information campaigns that there is a high rate of compliance in their industry or locality). [14]

6.0 Role of social partners in alert mechanisms

The involvement of social partners in alert mechanisms was given special attention by the participants. In some countries, the social partners are part of the assessment system. They inform the authorities, point to specific cases or suggest solutions for responsible authorities.

Ideas and ways of social partners' involvement in alert mechanisms discussed at the meeting:

▶ ELA representatives see a multiple-level based involvement of the social partners in the system of alert mechanisms;

▶ Social partners have valuable input to give and can form a vital part of the intergovernmental commission discussing access to data or data collection;

▶ Some social partners established hotlines and other complaint mechanisms for workers, whereby red-flag signals of undeclared work could be detected;

▶ The social partners can bring a personal perspective in creating alert mechanisms and contextualise the quantitate data analysis;

▶ Social partners should also be involved in designing the rules – algorithms in data mining;

▶ There is a need for better cooperation in the exchange of information/data between social partners, tax authorities and labour inspectors;

▶ Social partners can play a crucial part in alerting authorities about undeclared work and the violation of foreign workers' labour rights.

In Poland, the inspector must inform social partners about the scope, aim and result of inspections carried out. This process is important to interconnect the information. The inclusion of social partners is not only seen at the central level but also occurs at the local level. In Italy, the social partners are part of the inter-governmental commission, which meets twice yearly. The social partners requested, for example, to maintain the 'human side' in terms of the risks assessment, and to not only focus on the AI systems.

---

[14] Further reading: https://en.wikipedia.org/wiki/Scenario_planning;
https://ore.exeter.ac.uk/repository/bitstream/handle/10871/33035/JOTA-TEGreview-Malezieux.pdf?sequence=1&isAllowed=y .

7.0   Conclusions

7.1   Concluding remarks

The meeting participants discussed the necessary steps required to design an ideal alert mechanism in their country. They revealed their '**wish list'** to make progress in the creation of alert mechanisms.

One of the main challenges is ensuring access to data. To build effective alert mechanisms competent authorities within the Member States should have access to tax data. Cross-border data sharing, especially between labour inspectorates and other enforcement authorities would also be beneficial. There needs to be, however, a legal basis for this. While administrative databases have been created at various levels, the data are often fragmented and partially closed due to the legislation in place. Participants indicated that a more flexible legal system would enable access to information. The Platform has an important role to play in terms of sharing learning about what works in terms of accessing, using and sharing data.

Another challenge that hinders the development of alert mechanisms in some organisations is the shortage, or lack of data scientists. In this regard, more efforts are needed to increase institutional capacity and build competencies for relevant staff members. Participants concluded there was a need for specialised analysts or even a specialised unit to exploit data and create an interface with other authorities. Ideally, the income register should be universally accessible for all relevant authorities at national level.

An important element for developing alert mechanisms is the sharing of insights about the challenges encountered when conducting risk assessments. Participants also indicated that the social partners should play an important role in the development of alert mechanisms. Their industry knowledge, including information collected from workers, could be a valuable data source for identifying risks.

Participants suggested that collecting information at the EU level about all companies, such as size, number of employees, etc. would be very useful. The Internal Market Information System (IMI) is already in use, however the scope of the data exchange through IMI could be extended for the purposes of developing alert mechanisms.

Developing a centralised platform of data collection and coordination at the national level, similar to the system in Belgium for example, could help develop effective alert mechanisms. The inter-institutional cooperation in some Member States, however, needs to be improved, as in some cases institutions tend to compete rather than cooperate. Typically, data mining and data collection are divided among several institutions despite the fact the cases need to be assessed in one common platform.

Participants also pointed out the need to improve the data infrastructure at the national level (i.e., data digitalisation). To achieve this, political support is very important. However, it was noted that creating complex IT infrastructures is not enough, and that adequate human and financial resources are needed for system updates and maintenance. While political will is relevant, the need to produce consumer-friendly systems is even more critical. ELA can provide some practical guidance on data sharing and incorporate some aspects in Concerted and Joint Inspections (CJIs).

7.2   Key takeaways

Expand the alert mechanisms already in place

▶ Most of the participating enforcement authorities use some kind of alert mechanism to detect undeclared work at an early or later stage. Other countries can be inspired by, and build on, these experiences.

▶ The use of artificial intelligence and fiscal cash registers to design alert mechanisms are promising practices that are worth being explored in greater detail.

▶ Despite the fact that some of the alert mechanisms and risk assessments are applied at the micro level, they have the potential to be expanded and used in broader context.

Use multiple data sources

▶ Data used for alert mechanisms are mostly obtained from inspections records and employment registers, though other data sources are also important in this regard, such as hotline data, fiscal cash registers, and complaints registers.

▶ To predict undeclared work through data mining, there is a need for complex data integration and investigation procedures and predictors, based on data collected across institutions.

▶ To start small, less costly, qualitative forecasting methods could be used, such as future-basing.

Solutions to data sharing challenges

▶ Challenges to designing or using existing alert mechanisms stem from the legal challenges to connecting the databases, technical problems in building the datasets and data mining, involving the local level inspectors with their 'local' knowledge, and challenges to sharing the data as under the GDPR.

▶ The GDPR-based obstacles can be managed by an appropriate national legal framework. In many countries, data protection harmonisation is successful and allows data exchange.

▶ One solution to limited data access and exchange is the establishment of a culture of open data, and the creation of a space for safe data flow. The data flow works on demand, no data storing is needed, and the use of data is controlled.

▶ The lack of analysts and data scientists can be solved by establishing one central analytical unit - a centre that would serve for all such data requirements.

Exploit new technologies

▶ New technologies, such as artificial intelligence and machine learning, provide great opportunities to analyse large amounts of data, search for patterns and prevent undeclared work. The experience shows that the combination of algorithms and a 'human touch' ensures the best validity of the risk assessment tool.

▶ The high cost of artificial intelligence tools can be reduced by using pre-trained models and open public sources available to all; a good practice is to start small and continually expand the scope and datasets used.

Better collaboration across institutions and levels of operation

▶ Collaboration and involvement of various actors should be strengthened on a national and international level. Close cooperation agreements between labour inspectorates and/or other actors should implement ideas similar to Spain's business model and business rules.

- ▶ Moreover, social partners should also be involved at various levels in the system of alert mechanisms. Social partners have valuable input to give and can form part of the intergovernmental commissions or provide direct access to workers' experience with undeclared work which can serve as additional red-flag indicators.

- ▶ When creating alert mechanisms, a good practice is to involve the local inspectors and their knowledge as input to identify/verify the red-flag indicators. Therefore, it is essential to have close and trustworthy relations in place with local experts.

References

European Platform tackling undeclared work, (2017). Data Mining for More Efficient Enforcement. A Learning Resource Paper. Available at: https://ec.europa.eu/social/BlobServlet?docId=18746&langId=en

European Platform tackling undeclared work, (October 2017). 2017 Platform Survey report: organisational characteristics of enforcement bodies, measures adopted to tackle undeclared work, and the use of databases and digital tools. Available at: https://ec.europa.eu/social/BlobServlet?docId=18747&langId=en

European Platform tackling undeclared work, (September 2019). 2019 Annual Platform Survey: Tackling undeclared work in the collaborative economy and bogus self-employment, data exchange and protection, and cross-border sanctions. Available at: https://ec.europa.eu/social/BlobServlet?docId=21902&langId=en

Williams, C., Horodnic, I.A.(2020). Horizon scanning: early warning signals of future trends in undeclared work European Platform tackling undeclared work. Available at: https://www.ela.europa.eu/sites/default/files/2022-03/UDW_Study_report_on_early_warning_systems.2020_EN.pdf

Stefanov, R. (2021). Improving cross-government data and information exchange on national level and identifying good practices. Report of the peer learning dialogue. Platform tackling undeclared work. Available at : https://www.ela.europa.eu/sites/default/files/2022-01/PLD_Improving-cross-government-data-and-information-exchange.2021_EN.pdf