Co-funded by the
Erasmus+ Programme
of the European Union

InTheLoop4VET

O1-T3: Methodology for harvesting VET graduate data from open public data sources

Output type: Intellectual Output

PROMEA

May 2020

Co-funded by the Erasmus+ Programme of the European Union

Project acronym:	InTheLoop4VET
Project name:	Improving VET Providers' Uses of Graduate Tracking Feedback
Project code:	2019-1-SE01-KA202-060542

Document History

Versions	Date	Changes	Type of change	Delivered by
Version 1.0	07/01/2020	Initial document	-	PROMEA

Document Information

Document ID name:	2019-01-07_InTheLoop4VET_O1-T3_Methodology for harvesting VET graduate data from open public data sources
Document title:	O1-T3: Methodology for harvesting VET graduate data from open public data sources
Output type:	Intellectual Output
Date of delivery:	07/01/2020
Activity type:	Research/data compilation
Activity leader:	PROMEA
Dissemination level:	Public

Contents

1. Introduction

InTheLoop4VET is an Erasmus+ KA2 co-funded project that aims to address the need for Vocational Education and Training (VET) providers to strengthen their capacities for independent adaptation of their provision and quality determinants to labour market needs, through the establishment of graduate feedback loops and efficient graduate data handling and processing.

Vocational Education and Training aims to provide young people and workers with skills that ensure their employability and their access to further education. To what extend this objective is achieved is a key quality indicator of a VET system or of an individual VET provider. A better understanding of the performance of VET graduates in the labour market is also one of the key sources for assessing and improving the quality and labour market relevance of VET, alongside other methodologies such as forecasts of skills supply and demand[1]. Graduate tracking plays a key role in this context, as a principal mechanism for generating quantitative and qualitative data on graduates' employment and further training, skills-matching between training provision and job market requirements, as well as data on the actual efficiency and results of VET provision. The project aspires to contribute to the development of international standards in graduate tracking, by designing a transferable and customisable feedback loop mechanism aiming to secure a more efficient mobilisation of graduate tracking data in processes of VET adaptation at provider level.

The project will produce four Intellectual Outputs. The first output consists in the development of a graduate tracking toolkit that aims to increase the capacities of VET providers to generate useful data on graduates' career paths. Given the diversity of VET graduate tracking measures in terms of the level of implementation and type of tracking, this output develops three different methods of data generation and collection (survey, administrative, and synergistic). In particular, a survey tool and guides for data collection synergies and data harvesting from open data sources will be developed and delivered as an integrated tracking toolkit, while being usable in its separate functions as well. The toolkit will be designed for non-expert use by VET managerial and/or administrative staff and it will facilitate them to set up a data collection mechanism by taking advantage of three complementary ways of generating data on the progress of VET training recipients.

[1] *Mapping of VET graduate tracking measures in EU Member States Final Report*, Directorate-General for Employment, Social Affairs and Inclusion, Publications Office of the European Union, 2018

This document aims to provide non-expert/non-IT staff of VET institutions with guidelines on how to harvest VET graduate data from administrative and open public data sources. Specifically, the methodology will a) describe sources of data typically used for generating/extracting tracking data, b) propose steps to generating and/or extracting the data and c) designate a procedure for efficient sorting and storing of tracking data.

2. Open public data

2.1. What is open data?

According to the European Data Portal, Open Data is information that anyone can access, use and share with no limitations. There are two different aspects of openness: the legal and the technical. Open data should be technically open, in that the file is machine readable and non-proprietary where possible; in practice, this means that the data is free to access for everybody and the file format and its content are not restricted to a given non-open source software tool. The data should also be legally open, which in practice generally means that the data is published under an open licence and that the conditions for re-use are limited to attribution, permitting people to use the information in any way they want, including transforming, combining and sharing it with others, even commercially. In terms of costs, open data must be free to use, but this does not mean that it must be free to access. There is often a cost to creating, maintaining and publishing usable data. Ideally, any fee for accessing open data should be no more than the reasonable reproduction cost of the unit of data that is requested. This reproduction cost tends to be negligible for many datasets. Live data and big data can incur ongoing costs related to reliable service provision.

2.2. The benefits of open data

The overall benefits of open data are diverse and range from improved public sector performance, economic growth in the private sector to wider social welfare. Exploitation of open data can contribute to improving the quality and efficiency of public services and their processes, enabling cross-sector sharing of data, faster access to information and reduction of unnecessary spending. Social welfare can be improved through the enhancement of collaboration, active participation of citizens and social innovation, as well as through the promotion of transparency, democratic control and accountability.

2.2.1. The benefits of open data for VET graduate tracking

The exploitation of publicly available datasets is an efficient methodological approach of supplementing data gathered via other means, such as ad-hoc surveys, enabling VET providers to have a more comprehensive view of their graduates' destinations, and to compare their results with the results of other education providers (VET or non-VET) at regional or national level. The benefits of using open data are multiple:

- Open data provide the possibility to compare different groups of graduates, gain better insights into the functioning of the education system and compare the progression of graduates from different educational pathways and levels.
- Open data, for example from publicly available databases of statistical authorities, can provide factual information, subject to less errors, which can add value to the qualitative insights gathered via surveys.
- Open data enable VET providers to identify and confirm trends in the trajectories of graduates (for example the results regarding the employment status of graduates measured through a sample-based survey can be contrasted with the data on employment status from an open database covering the whole target population).
- The costs of collecting information can be significantly reduced, in that the use of available datasets reduces the need to design and run surveys.
- Open data can help fill information gaps when surveys have low response rates (especially when it comes to target groups that often suffer from survey fatigue/over-surveying).
- Making use of open datasets can help make surveys shorter by limiting the number of questions on factual information (which would already be available on open databases) and thus increase response rates.

2.3. Open data policy in the EU

Over the last few years, an open data ecosystem has been developing globally. Led by the United Kingdom and the United States, many countries, regions, and local authorities have started to make their data broadly available via portals and websites for any kind of use the public wants to make of it. At the level of the European Union, the European Commission also adopted an Open Data Strategy in December 2011, aiming to make open data the standard in the Member States of the European Union (European Commission, 2011).

While the European Commission seems only recently to have picked up the topic of open data, it has actually been striving for a greater availability of data held by the public sector for over 20 years, by means of its policy on the re-use of public sector information. In 2003 this led to the adoption of the Directive on the re-use of public sector information (PSI Directive), which was revised in 2013[2]. Public Sector Information (PSI) is the wide range of information that public-sector bodies collect, produce, reproduce, and disseminate in many areas of activity while accomplishing their institutional

[2] Janssen, K. and Hugelier S., "Open data as the standard for Europe? A critical analysis of the European Commission's proposal to amend the PSI Directive", European Journal of Law and Technology, Vol. 4, No. 3, 2013

tasks. PSI may include (among others) social, economic, geographical, cadastral, weather, tourist, and business information. If PSI is made available under an open licence, it is called Open Government Data. The general term Open Data also refers to other types of non-public-sector data that is freely available, for example social media data[3].

Since 2015, the European Data Portal (https://www.europeandataportal.eu/en/) has been monitoring the development of national Open Data policies and Open Data portals throughout Europe and has provided an assessment of the Open Data maturity level across the EU Member States as well as Liechtenstein, Norway and Switzerland – referred to as the EU28+ on an annual basis, including Iceland since 2017. The assessment measures the presence of an Open Data policy, the use of data and the political, social and economic impact of Open Data within European countries.

According to the 2018 assessment report[4] all EU28 countries have a dedicated Open Data policy in place. In 2018, 23 EU Member States were successful in enabling the development of Open Data initiatives at local or regional level. 15 Member States (56%) stated in 2018 that all their published Open Data can be accessed free of charge (Austria, Croatia, Cyprus, Czech Republic, Denmark Estonia, Finland, France, Italy, Latvia, Luxembourg, Netherlands, Poland, Portugal, and Slovenia). In another 9 EU countries amounting to 33% (Belgium, Bulgaria, Greece, Ireland, Lithuania, Romania, Slovakia, Spain and Sweden) 90-99% of data can be accessed free of charge. In the UK and Germany this percentage is slightly lower and ranges between 75% and 89%. In 2018, the most popular data domain was Government and Public Sector with 58% of countries mentioning it in their top 5, followed by Population and Social Conditions (50%), Energy and Environment (46%), Transport Infrastructure (42%), Economy and Finance (42%), Regions and Cities (35%), Education and Culture (35%), Agriculture, Fisheries, Forestry & Foods (12%), Justice, Legal System, Public Safety (12%), Social Mobility and Welfare (8%), Health (8%), Housing (4%), and Statistics (4%).

2.4. Open data policy in Sweden

The Government Offices of Sweden are working to make data accessible and to promote greater openness and better service in the public sector. Public authorities are expected to make their public data accessible for reuse, free of charge or on standardised and generous terms. Public data about

[3] Wendy Carrara, Sem Enzerink, Fréderique Oudkerk, Cosmina Radu, Eva van Steenbergen (Capgemini Consulting), "Open Data Goldbook for Data Managers and Data Holders – Practical guidebook for organisations wanting to publish Open Data", European Data Portal, 2018

[4] https://www.europeandataportal.eu/sites/default/files/edp_landscaping_insight_report_n4_2018.pdf

the Government's work is currently available on the Riksdag's website, since most of the documents (legal documents) the Government sends to Riksdag are available via the Riksdag's open data.

A general process for making datasets available is implemented by the Swedish Environmental Protection Agency. This is a translation of a generic process for making datasets available to different target groups. It was initially intended to be used for open data but is applicable for all types of data publication both within an organisation and to other external parties. Some of the content might be specific to Sweden (i.e. regulations for information security) but is understandable for everyone in most parts.

Also, Swedish universities are authorities and the research organisation's records and documents are often official and may therefore be considered public (the principle of public access to official records) when secrecy does not apply. This concept is defined in 2 kap. 3 § tryckfrihetsförordningen (Chapter 2, 3§ of the Freedom of the Press Act). Documents that are public ("allmäna") must be disclosed to those asking to see it. Only with reference to a rule in the secrecy act can one decide that they should not be considered official ("offentliga") documents, but rather as secret.

2.5. Open data policy in Greece

The Greek regulatory framework regarding access to public data was mainly established through the adoption of European Directives, particularly the PSI directive on access to public sector information.

There are several laws and government bulletins that constitute the Greek regulatory framework on open data, the most important of which are:

- Law No. 4305/2014 (Government Gazette 237/A) "Open disposal and re-use of public sector documents, information and data, amendment of Law 3448/2006 (A 57), adaptation of national legislation to the provisions of the Directive 2013/37/EU of the European Parliament and of the Council, further enhancing transparency, Regulations of the ECHR Introductory Contest and other provisions".
- Law No. 3448/2006 (Government Gazette 57/A) "On the re-use of public sector information and the regulation of matters of competence of the Ministry of Internal Affairs, Public Administration and Decentralization".

2.5.1. Open government initiatives

Data.gov.gr

Data.gov.gr is an open data policy implementation tool enabling the enforcement of the relevant legislation, following the integration of Directive 2013/37/EU of the European Parliament and of the Council of 26 June 2013 amending Directive 2003/98/EC on the re-use of public sector information.

Data.gov.gr is the central directory of public data that provides access to databases of Greek government agencies. The purpose of data.gov.gr is to increase access to high-value, readable datasets by providing unified cataloguing, indexing, storage, retrieval and availability of public sector data and information, as well as online services to citizens and third party information systems.

Diavgeia

Diavgeia (Greek word for clarity) or "Transparency Program" is an initiative that began October 1st 2010 in Greece, according to which all government institutions are obliged to upload their acts and decisions on the Internet with special attention to issues of national security and sensitive personal data. Each document is digitally signed and assigned a unique Internet Uploading Number (IUN) certifying that the decision has been uploaded at the "Transparency Portal". Following the latest legislative initiative (Law 4210/2013) of the Ministry of Administrative Reform and e-Governance, administrative acts and decisions are not valid unless published online.

The main objectives of the Transparency Program concern: a) safeguarding transparency of government actions, b) eliminating corruption by exposing it more easily when it takes place, c) observing legality and good administration, d) reinforcing citizens' constitutional rights, such as the participation in the Information Society, e) enhancing and modernizing existing publication systems of administrative acts and decisions, and f) making of all administrative acts available in formats that are easy to access, navigate and comprehend, regardless of the citizens' knowledge level of the inner processes of the administration. Public authorities adopted the Program in three phases: Ministries in October 2010, Extended Public Sector and Independent Authorities in November 2010, Regional and Local Authorities in March 2011.

Geodata.gov.gr

Geoadata.gov.gr is providing open geospatial data and services for Greece, serving as a national open data catalogue, an INSPIRE-conformant Spatial Data Infrastructure, as well as a powerful foundation for enabling value added services from open data. Operating since 2010, geodata.gov.gr was one of the first open data catalogues in the world, contributing to the national and international

open government agenda. The data are offered with the specific terms of Greek Creative Commons license Attribution (CC BY v.3.0) as specified in detail in the relevant License Use. This means that one can use them any way they want, improve them, edit them, create commercial applications, or anything else needed, indicating the source of origin, i.e. geodata.gov.gr.

2.6. Open data policy in Spain

With the approval by the European Parliament of Directive 2003/98/EC, a common framework of reference on the reuse of public information was established for the Member States of the EU, who incorporated this regulation into their respective legal systems and developed a concrete political framework for each country. Applicable in Spain, the most important regulations that form the open data legislation include:

- Law 37/2007 of November 16 on the reuse of public sector information,
- Royal Decree 1495/2011 of October 24, which develops Law 37/2007 on the reuse of PSI for the scope of the state public sector,
- Directive 2013/37/EU of the European Parliament and of the Council of June 26 2013 amending Directive 2003/98/EC on the reuse of PSI,
- Law 18/2015 of July 9 amending Law 37/2007 on the reuse of PSI.

2.6.1. Open government initiatives

Aporta Initiative

It is an open data initiative of the Government of Spain. The Aporta Initiative was launched in 2009 to promote the opening of public information and development of advanced services based on data. It is backed by the Ministry of Economy and Business, the Ministry of Territorial Policy and Civil Service and the Public Corporate Entity Red.es. Developed under the current legislative framework, it includes seven lines of action, which are reflected on the datos.gob.es, the platform used as a meeting point for administrations, businesses and citizens who are part of the open data ecosystem in Spain (https://datos.gob.es/es).

The main goal of the Aporta Initiative, a key element in the Spanish government's open data policy, is to harmonize and efficiently take advantage of the synergies between ongoing open data projects. It seeks to always drive and coordinate actions being carried out by different levels of the administration, the private sector and academic field, according to an integrating governance model. It does all of this in order to promote new products and services from the private sector and civil society to benefit society. The Aporta Initiative and datos.gob.es are designed for all data ecosystem

agents: users, citizens or professionals who demand data and/or want to find out about the latest news, application or services related to data, public entities who provide and use public data, and who want to be up-to-date with sector news, reusers and infomediaries who require data sources to create products and services they want to publicise.

Valencian Community

In accordance with article 4.1 of Law 2/2015, of April 2, on Transparency, Good Governance and Citizen Participation of the Valencian Community, public information is understood as the set of contents or documents that are in the possession of any of the subjects included in the scope of application of this law and that have been developed or acquired in the exercise of their functions.

Any citizen, individually or in representation of any legally constituted organization, has the right of access to public information. It will be done by prior request and without more limitations than those provided by law.

2.7. Open data policy in Lithuania

The most important directives and laws that constitute the regulatory framework on open data in Lithuania are the following:

- The Civil Code of the Republic of Lithuania
- The law on the Management of State Information Resources of the Republic of Lithuania
- The law of the Republic of Lithuania on the *right to receive information from state and municipal institutions and bodies*
- The Resolution of the Government of the Republic of Lithuania "on the approval of the description of the procedure for the establishment, creation, reorganization and liquidation of the registers"
- Recommendations on the opening of public sector data, approved by the Minister of Transport and Communications
- Recommendations for data submission formats and standards approved by the Director of the Information Society Development Committee under the Ministry of Transport and Communications.

2.7.1. Open government initiatives

The open data portal (https://data.gov.lt/) with its all functions will be applied by June 2020, even though it was started in January 2020. The open data portal is being created as unified linking point to all Lithuanian open data sets. This open data portal was mainly established by adopting European

Directives based on open data. Its focus is to provide technological means for open data suppliers to prepare and publicize metadata of open data sets. Furthermore, it will provide an opportunity for the open data users to complete a search of necessary open data sets and will give a link to the sets of open data.

The purpose of creation of the open data portal includes several areas:

- Public sector institutions and business, which will be able to:
  - Complete the checking of the processed open data sets,
  - Plan open data completion,
  - Form metadata of datasets,
  - Use of techniques of data transporting and form open data sets for publication and usage.
- Users of open data will be able to:
  - Easily and with comfort find necessary data sets,
  - Give remarks for the specification and complementation of the publicized data sets,
  - Register the potential open data demand,
  - Share the innovative decisions of data application and analysis as well as give other suggestions.

The main principles of open data policy are:

- **Availability and Access:** the data must be available as a whole and at no more than a reasonable reproduction cost, preferably by downloading over the internet. The data must also be available in a convenient and modifiable form.
- **Re-use and Redistribution:** the data must be provided under terms that permit re-use and redistribution including the intermixing with other datasets.

Universal Participation: everyone must be able to use, re-use and redistribute - there should be no discrimination against fields of endeavour or against persons or groups. For example, 'non-commercial' restrictions that would prevent 'commercial' use, or restrictions of use for certain purposes (e.g. only in education), are not allowed.

2.8. Open data policy in Germany

The most important acts that formed the regulatory framework on open data in Germany are the following:

- The Law on the re-use of Public Sector Information came into force on 19 December 2006, transposing the pertinent EU Directive 2003/98/EC. The law regulated the information re-usage

of public sector information beyond the public-administrative scope, namely, for shaping new information-products and related services. The law specifids that re-usage of public sector information had to be non-discriminative, contemporary and not exclusive. In accordance with the revision of the EU Directive, the law was revised in July 2015.

- After being narrowly approved by the Lower Chamber of the Bundestag in early June 2005, Germany's Freedom of information (FOI) Act was voted on 8 July 2005 by the Bundesrat, published in the Federal Gazette on 13 September 2005 and came into force on 1 January 2006. The law provided the public with a general right to access Federal Government information. However, this general right was limited by a number of broadly defined exemptions, covering, for instance, security-sensitive issues, potential threats to public safety and even the 'fiscal interests of the Federal Government'. The legislation also contained an 'Internet clause' to compel federal administration bodies to make a number of items publicly available online. Certain Federal States also have their own Freedom of Information Legislation.
- On 17 June 2013, the Bundestag, with the approval of the Bundesrat, adopted the Act to promote electronic government (eGovernment Act, EGovG). The German law for the promotion of eGovernment came into effect on 1 August 2013. Its aim was to facilitate electronic communication with the administration and to enable federal, state and local authorities to provide simpler, more user-friendly and efficient eGovernment services. The main provisions included:
  - Obligation for federal, state and local authorities to open a point of access for the transfer of electronic documents.
  - Obligation for federal authorities to open a De-Mail access and to offer an electronic proof of identity via electronic identity (eID) card and electronic residence permit.
  - Electronic record keeping.
  - Enabling electronic evidence and electronic payment in administrative procedures.
  - Obligation to document and analyse processes before implementing information technology systems.
  - Provision of machine-readable data (open data).
  - Georeferencing of electronic registers.
  - The eGovernment Act triggered legislative activities in most of the German federal states. By now seven *Länder* have already adopted their own eGovernment Act while, several others have initiated legislative procedures.

- On 18 May 2017, the German Bundestag passed the draft of Federal Open Data Act amending the eGovernment Act, presented by the Federal Minister of the Interior, which initiates cultural change in the administrations and obliges federal authorities to publish unprocessed data in the future. The law sets out common criteria for open data. These include in particular the free of charge provision, the free access to the data, as well as the machine readability. At the same time, data protection, as well as other reasons that prevent publication are observed, thus ensuring that only data suitable for publication is provided. So that the open data provided by the different administrations can be easily found, the nation-wide metadata portal GovData was created.

2.8.1. Open government initiatives

On 20-21 September 2019, the Federal Ministry of the Interior, Building and Community (BMI) launched an initiative for Regional Open Government Laboratories. The initiative will establish up to 16 laboratories to test new collaborations between administrative bodies, politicians, citizens and social interest groups. In addition, this initiative aims to encourage people to participate in regional development and to facilitate a more open discussion and debate between local politics, citizens, external partners and interested stakeholders. The launch took place during the 2nd edition of the Creative Bureaucracy Festival in Berlin, Germany. The event was attended by innovators from all levels of the Public Sector in Germany and abroad, including academics, government ministers and representatives from municipalities and international organisations.

2.9. Types of published data

Data is created, stored, and distributed covering a large variety of topics and categories. However, not all types of data are of equal relevance or value for citizens. In 2013 the G8[5] came together to discuss governmental transparency, innovation and accountability. This discussion led to the creation of the "G8 Open Data Charter": a summary of visions and principles for creating a transparent government, the opening up of data and its quality and quantity. According to this charter, the G8 High Value categories of data are summarized in the following categories:

Data category	Examples of datasets
Companies	Company or business register

[5] The Group of Eight (G8) refers to the group of eight highly industrialized nations—France, Germany, Italy, the United Kingdom, Japan, the United States, Canada, and Russia—that hold an annual meeting to foster consensus on global issues like economic growth and crisis management, global security, energy, and terrorism.

Crime and justice	Crime statistics, safety
Earth observation	Meteorological/weather, agriculture, forestry, fishing, and hunting
Education	List of schools, performance of schools, digital skills
Energy and environment	Pollution levels, energy consumption
Finance and contracts	Transaction spend, contracts let, call for tender, future tenders, local budget, national budget (planned and spent)
Geospatial	Topography, postcodes, national maps, local maps
Global development	Aid, food security, extractives, land
Government accountability and democracy	Government contact points, election results, legislation and statutes, salaries (pay scales), hospitality/gifts
Health	Prescription data, performance data
Science and research	Genome data, research and educational activity, experiment results
Statistics	National statistics, census, infrastructure, wealth, skills
Social mobility and welfare	Housing, health insurance and unemployment benefits
Transport and infrastructure	Public transport timetables, access points broadband penetration

The purpose of the above list of categories is to ensure that data holders focus on the release of the right and most relevant types of data. This does not mean that other categories of data cannot be published. The list above gives an indication of the topics that should have the highest priority, as these datasets are indicated as datasets with the highest potential value. Based on the G8 Open Data Charter and different European Commission studies, the European Commission published its own standards to help harmonise European efforts related to licences and datasets and accommodate the PSI Directive as a result.

2.10. Categorization of published data

Given the vast amount of published data, it is important to navigate through the right categories, when searching for information on a specific sector. Data categories are categorisations of datasets linked to a common theme. Each data portal usually creates its own set of data categories. An example of categorisation is the one implemented by the European Data Portal. The European Data portal harvests the metadata of Public Sector Information available on portals, national, regional and local, throughout Europe. Currently there are 71 catalogues which are harvested, with a total

amount of over 616,000 datasets available on the portal. These datasets are divided in the categories shown in the following figure:

Figure 1

3. Sources of relevant open public data

3.1. Existing sources used for generating/extracting tracking data in EU

VET graduate tracking measures are defined as tools which collect information for analysis of graduates' educational achievements on completion of VET, their further education and higher achievements, and their employment. According to findings of a study[6] on graduate tracking measures conducted by the European Commission, VET graduate tracking measures vary among EU Member States; some countries do not have any measures, others do not have regular measures, and many of them use either administrative data or survey data. In total, 24 EU Member States have VET graduate tracking measures, of which 19 have national measures. Regarding the InTheLoop4VET partnership countries, Spain, Germany and Sweden belong to those EU countries that have measures, which include employment and education indicators. Greece belongs to the countries that do not have VET graduate tracking measures.

All EU Member States regularly collect a variety of information (register/administrative data) about their citizens that could potentially be used for VET graduate tracking. This includes data compiled in registers concerning education, unemployment, social security, taxes, population, ESF beneficiaries, and other activities such as housing, enterprises/business, pensions, employment and contracts. Of the various tracking measures that have been identified (85 in total), 59 used surveys and 35 used administrative learning and employment data such as social security and tax data. Germany and Sweden use administrative and survey data. Longitudinal measures are rarely implemented (only one country has currently such measures in place). Nine countries collect information on graduate results up to one year after completing their studies, but not later. Moreover, there are currently no EU-level VET graduate tracking measures; some one-off tracking measures were developed in the past, most focused on higher education graduates (CHEERS, REFLEX, HEGESCO). One EU-level study examined the scope to compare national tracking measures of secondary education – including VET graduates (CATEWE).

While no European-wide VET graduate survey has been developed, there are EU-level cross-sectional surveys designed for other purposes that could enable the measurement of graduates' progression into the labour market: the EU Labour Force Survey (EU-LFS) and its ad hoc modules on young people in the labour market, the Survey on Income and Living Conditions (SILC), the Adult Education Survey (AES) and the OECD Programme for the International Assessment of Adult

[6] *Mapping of VET graduate tracking measures in EU Member States Final Report,* Directorate-General for Employment, Social Affairs and Inclusion, Publications Office of the European Union, 2018

Competencies (PIAAC) survey. These could potentially be useful to track VET graduates as they all differentiate between general and vocational education. Also, the SILC and EU-LFS collect some longitudinal data which could be used for tracking VET graduates.

To be of use to the analysis of VET graduates' transition into the labour market, the international and national surveys have to provide a sufficient sample of VET graduates for analysis at EU and national level (and also for comparison between Member States), and to differentiate between VET and general education graduates. Large-scale surveys meet these conditions to a higher extent than smaller-scale surveys, since they have wider overall samples and consequently larger samples of VET graduates.

The EU Labour Force Survey

The European Union Labour Force Survey (EU-LFS) is conducted in all Member States of the European Union, 4 candidate countries and 3 countries of the European Free Trade Association (EFTA) in accordance with the Council Regulation (EEC) No. 577/98 of 9 March 1998. The EU-LFS is a large household sample survey providing quarterly results on labour participation of people aged 15 and over as well as on persons outside the labour force. All definitions apply to persons aged 15 years and over living in private households. Persons carrying out obligatory military or community service are not included in the target group of the survey, as is also the case for persons in institutions/collective households.

The data collection covers the years from 1983 onwards. In general, data for individual countries are available depending on their accession date. The Labour Force Surveys are conducted by the national statistical institutes across Europe and are centrally processed by Eurostat. The national statistical institutes are responsible for selecting the sample, preparing the questionnaires, conducting the direct interviews among households, and forwarding the results to Eurostat in accordance with the requirements of the regulation. In 2018, the quarterly LFS sample size across the EU was about 1.5 millions of individuals. The EU-LFS covers all industries and occupations. A significant amount of data from the survey is available in Eurostat's online dissemination database, which is regularly updated and available free of charge. Access to microdata can only be granted to EUI researchers. The EU-LFS is the main data source for the domain "employment and unemployment" in the database. The contents of this domain include tables on population, employment, working time, permanency of the job, professional status etc. The data is commonly broken down by age, sex, education level, economic activity and occupation where applicable.

Adult Education Survey

The Adult Education Survey (AES) covers adults' participation in education and training and is one of the main data sources for EU lifelong learning statistics. The AES covers the resident population aged 25-64, providing information on: a) participation in formal education, non-formal education and training and informal learning, b) characteristics of the learning activities, c) volume of instruction hours, d) reasons for and obstacles to participating, e) access to information on learning possibilities, f) employer financing and costs of learning, and g) self-reported language skills.

Adult Education Surveys were carried out in 2007, 2011 and 2016 and results are published in Eurostat's online database. The next AES is planned for 2022.

Statistics on income and living conditions

The European Union Statistics on Income and Living Conditions (EU-SILC) aims at collecting timely and comparable cross-sectional and longitudinal multidimensional microdata on income, poverty, social exclusion and living conditions. The EU-SILC project was launched in 2003 in six Member States (Belgium, Denmark, Greece, Ireland, Luxembourg and Austria) and Norway. The EU-SILC legal basis entered into force in 2004 and covers now all EU countries, Iceland, Norway, and Switzerland. Data has been collected in a consistent manner since 2005; in view of this, the data can be considered to be comparable and reconcilable over time.

The reference population in EU-SILC includes all private households and their current members residing in the territory of the countries at the time of data collection. All household members are surveyed, but only those aged 16 and more are interviewed. EU-SILC results are available in Eurostat website and comprise multi-dimensional datasets and policy indicators. Both datasets and indicators are update on Eurostat website as soon as new data become available. For scientific purposes only, access to anonymized microdata is also possible under specific conditions.

Small-scale surveys

Smaller-scale international surveys such as the Eurofound Survey on Working Conditions, Cedefop Opinion survey on VET in Europe and Cedefop European skills and jobs (ESJ) survey collect data on EU VET graduates' current situation in the labour market. The most relevant data collected through these surveys concern – among others – the highest level of education attained, difficulties finding a job, time elapsed until first long-term job after finishing studies, satisfaction with professional career, further education, gross monthly earnings, match between the respondent's skill and those required in their job, and type of employment contract.

3.2. Sources of data in Sweden

Statistics Sweden

Statistics Sweden (SCB) is the Swedish government agency responsible for developing, producing and disseminating official national statistics. SCB's database includes a considerable amount of official statistics in different areas and is available on the website (https://www.scb.se/en/finding-statistics/). Specifically, it provides access to data related to, among other subjects, population (size, average life expectancy, forecasts, immigration/emigration etc), social insurance (disability, families with children, pensions, etc), social services (care services, drug abuse, economic support etc), labour market (employment, forecasts, unemployment, work environment, wages etc), living conditions (gender equality, families, ICT, integration, leisure etc), household finances (expenditure, housing, incomes, tax etc), judicial system (offences, persons convicted of offences, the correctional care system), and education and research (grades, financial aid, higher education, school, staff etc).

Swedish National Data Service

The Swedish National Data Service (https://snd.gu.se/en) has a primary function to support the accessibility, preservation, and re-use of research data and related materials. Together with a network of more than 30 universities and public research institutes, they strive to create a national infrastructure for open access to research data.

LADOK

LADOK is a student administration system used in all Swedish universities and college universities. The acronym "Ladok" abbreviates the Swedish "Lokalt adb–baserat dokumentationssystem" ("Local Edp based documentation system"). It is a student registration and grading documentation system.

LADOK has been developed by the universities and has been controlled by Ladokkonsortiet ("the Ladok Consortium"). Each institution using the system has its own database which is limited to containing the results and registrations of its students; each handles its own service management on one of three certified service nodes, which are located in Umeå, Uppsala, and Lund. LADOK has combined data and designed a database that allows them to follow student from before, at entry, during and after at institutional level and programme level.

Swedish National Agency for Higher Education

The Swedish National Agency for Higher Education is a Government agency in Sweden that oversees the Swedish public-school system for children and adults. This agency is responsible for official statistics in the area of education.

The Swedish National Agency for Higher Education monitors and analyses developments and trends within Swedish higher education. They are also responsible for official statistics on higher education. The statistics they collect include, for instance, figures on the number of students in different programmes, tracking students after graduation, distribution of the teaching staff by age and gender, as well as the higher education institution's financial reporting.

3.3. Sources of data in Greece

Hellenic Statistical Authority

The Hellenic Statistical Authority (ELSTAT, website: https://www.statistics.gr/en/home/) is the national statistical service of Greece. It cooperates with public and private agencies in Greece or abroad, such as educational institutions, research centers and non-profit organisations for the promotion of scientific research for statistical issues, providing data across sectors including economy, population and social conditions, agriculture, industry, trade, transport, environment and energy. Specifically, the Hellenic Statistical System comprises agencies that have the responsibility or obligation to collect statistical data; the role of ELSTAT is to coordinate all the activities of the other agencies of the system that concern the development, production and dissemination of the country's official statistics and forwards these statistics to Eurostat. The services and agencies of the public sector, the Legal Entities under Private Law, the associations of individuals and natural persons are obliged to grant ELSTAT access to all the administrative sources, public registers and files they keep, in printed, electronic or other form, and provide, in an accurate and timely manner, ELSTAT with primary statistical data and information, which is required for the performance of its duties.

3.4. Sources of data in Spain

National Statistics Institute

The National Statistics Institute (Spanish: Instituto Nacional de Estadística, INE) is the official agency in Spain that collects statistics about demography, economy, and Spanish society. It is an autonomous organisation responsible for overall coordination of statistical services of the General

State Administration in monitoring, control and supervision of technical procedures. Every 10 years, this organisation conducts a national census (the last census took place in 2011). All the updates of the different fields of study can be accessed at INEbase, the system the INE uses to store statistical information: https://www.ine.es/dyngs//INEbase/en/listaoperaciones.htm. The information on the database of INE is organised in a way that makes it possible to ascertain all information available for each topic: operations for which results are presented, together with a short description of the variables published, the intervals and availability of data and the geographical scope; publications and related studies; links to other websites in order to broaden information from external sources; and a link to the Inventory of Statistical Operations of the State General Administration, in order to ascertain all Spanish Statistical System operations relating to the topic.

3.5. Sources of data in Lithuania

Statistics Lithuania

The Lithuanian Department of Statistics (Statistics Lithuania) is a public authority that participates in developing and implementing public policy in the field of statistics assigned to the Minister of Finance, and coordinating official statistics in the country pursuant to the provisions of the Official Statistics Programme[7]. Statistics Lithuania applies statistical methods for the collection, processing, analysis of statistical data and releases statistical information on economic, social, demographic, and environmental changes in the country at the national and regional levels. Information prepared according to the Law on Official Statistics is available and provided free of charge to everyone interested in the country's economic and societal development. The database of Statistics Lithuania can be accessed at the Official Statistics Portal: https://osp.stat.gov.lt/web/guest/pradinis.

Open data portal

The Open Data portal (https://data.gov.lt/) is a database where the data from various institutions are stored: National Education agency, National Paying Agency, Center of Agricultural Information and Rural Business, Employment services, State Labor Inspectorate. This database stores information of various data sets and their metadata. This portal is going to include information and data from other portals such as Spatial Information Portal of Lithuania (https://www.geoportal.lt/geoportal/), Open Finance Information System of The Ministry of Finance, and other.

[7] Official Statistics Programme means the annual list of the statistical surveys to be conducted by the bodies managing official statistics for public demands and for the needs of public administration and commissioned by legal acts of the EU, laws of the Republic of Lithuania, resolutions of the Government of the Republic of Lithuania and resolutions of the Board of the Bank of Lithuania.

Register of Institutions of Education and Higher Education

The National Agency for Education (https://www.nsa.smm.lt/en/infrastrukturos-pletros-departamentas/registrai-pagrindinis-puslapis/svietimo-mokslo-instituciju-registras/) is responsible for the register of information about educational institutions (higher education, secondary, informal educational institutions, VET institutions, institutions of educational support), scientific research institutes, their departments, departments of foreign higher educational institutions founded in Lithuania. The main task of the register is to register the objects, collect, process, store, systemize, save and provide data to physical and legal persons.

3.6. Sources of data in Germany

Federal Statistical Office

The Federal Statistical Office is a federal authority of Germany reporting to the Federal Ministry of the Interior. The Office is responsible for collecting, processing, presenting and analysing statistical information concerning the topics economy, society and environment. Its database, GENESIS, (https://www-genesis.destatis.de/genesis/online) provides access to tables, graphs and press releases with key statistical information on the following topics:

- Society and environment (population, education, research and culture, income, consumption and living conditions, health, social statistics, environment, housing etc)
- Economy (foreign trade, prices, national accounts, domestic product etc)
- Economic sectors and enterprises (construction, services, energy, transport, agriculture and forestry, fisheries etc)
- Labour (labour market, labour costs, non-wage costs, earnings)
- Government (justice, public finance, taxes, etc)
- Countries and regions (regional statistics, Europe, international statistics).

Statistical offices of the German states

There are 14 statistical offices in Germany for the 16 states, which are responsible for carrying out the task of collecting official statistics together and in cooperation with the Federal Statistical Office. The implementation of statistics according to the German constitution is executed at state level and the data can also be accessed at the websites of the respective statistical institutions of Baden-Württemberg, Bavaria, Berlin and Brandenburg, Bremen, Hamburg and Schleswig-Holstein, Hesse, Lower Saxony, Mecklenburg-Vorpommern, North Rhine-Westphalia, Rhineland-Palatinate, Saarland, Saxony, Saxony-Anhalt, and Thuringia.

All datasets of the Federal Statistical Office and the statistical offices of the German states can also be accessed through the Joint Statistical Portal (Statistische Ämter des Bundes und der Länder Gemeinsames Statistikportal): http://www.statistikportal.de/de, which is developed and operated by the federal and state statistical offices.

Federal Employment Agency

As part of official statistics in Germany, the Federal Employment Agency (Bundesagentur für Arbeit) creates and publishes statistics on the labour market and basic security for employment seekers for all regions. The labour market and basic security statistics are compiled as departmental statistics under the specialist supervision of the Federal Ministry of Labour and Social Affairs and with a high degree of topicality in order to give the institutions that are directly involved in the labour market a reliable basis for assessing the overall situation and regional developments. The Agency's statistical database (https://statistik.arbeitsagentur.de) provides key figures on the German labour market, which are updated monthly, as well as a report on the European labour market situation. It also provides reports and tables on different topics and geographic areas as well as specialized reports and interactive visualizations on different subjects.

4. Extracting comparable data about VET graduates

The main purpose of this methodology is to provide VET institutions with recommendations on how to find public data relevant to VET graduates and use them mainly for benchmarking, in order to be able to assess the quality and effectiveness of their provision compared to national and European standards and aggregated results. Therefore, the methodology will focus on datasets that provide direct information about vocational education, rather than non-self-explanatory datasets that would indirectly lead to potential conclusions about VET quality.

The steps a VET provider is advised to follow, in order to find useful data in the context of deploying the InTheLoop4VET toolkit, will be elaborated using the database of Eurostat as an example. National statistical authorities apply similar categorization to their databases, so the following process is transferable.

4.1. Step 1 – find the relevant categories

Depending on the categorization each source uses, information about VET can be found in multiple data categories such as “population and social conditions”, “economy and finance”, “industry, trade and services”, and “science and technology”. Therefore, filtering with appropriate keywords is very important in terms of research efficiency. A simple search typing the keyword “vocational” in the search engine of Eurostat’s database produces 881 results[8], which include datasets, statistical books, statistical reports, news releases, manuals and guidelines across thematic categories (e.g. economy and finance, transport, science and technology, general and regional statistics etc). Navigating through so many results would require a lot of time, so it is necessary to narrow these down with more filters.

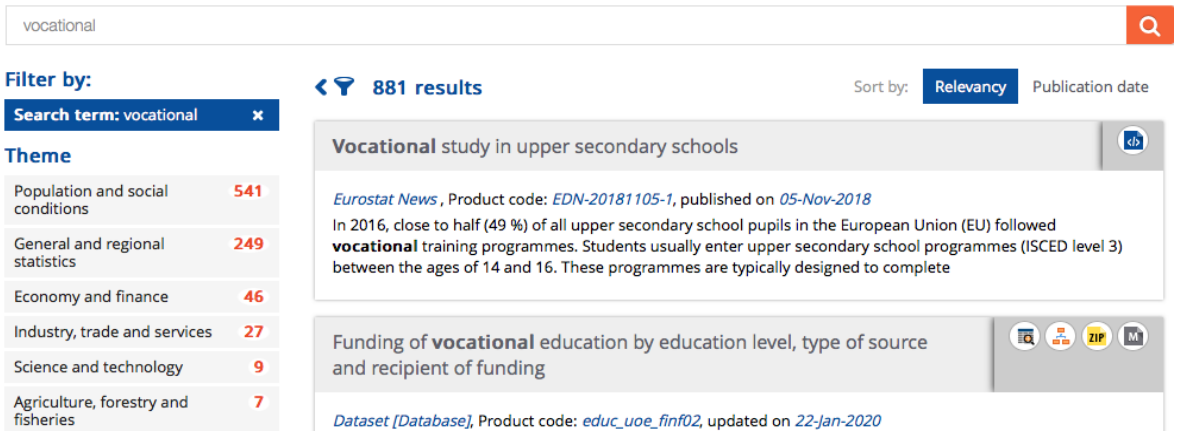

Figure 2

[8] The database of Eurostat is constantly being updated, so the number of results varies with time.

4.2. Step 2 – narrow down the results

The thematic categories that produce the most results in Eurostat statistics are “population and social conditions” (541 results) and “general and regional statistics” (249 results). These are also the categories, which are more likely to produce results that are more relevant and useful in deploying the graduate tracking toolkit for VET providers. Therefore, it is recommended to apply further filters by selecting one of the aforementioned categories.

Filtering for example by theme “population and social conditions” reduces the total number of results to 541 (Figure 3); these however include many old datasets with publication dates starting from 1997, which would be outdated in the context of tracking current graduates.

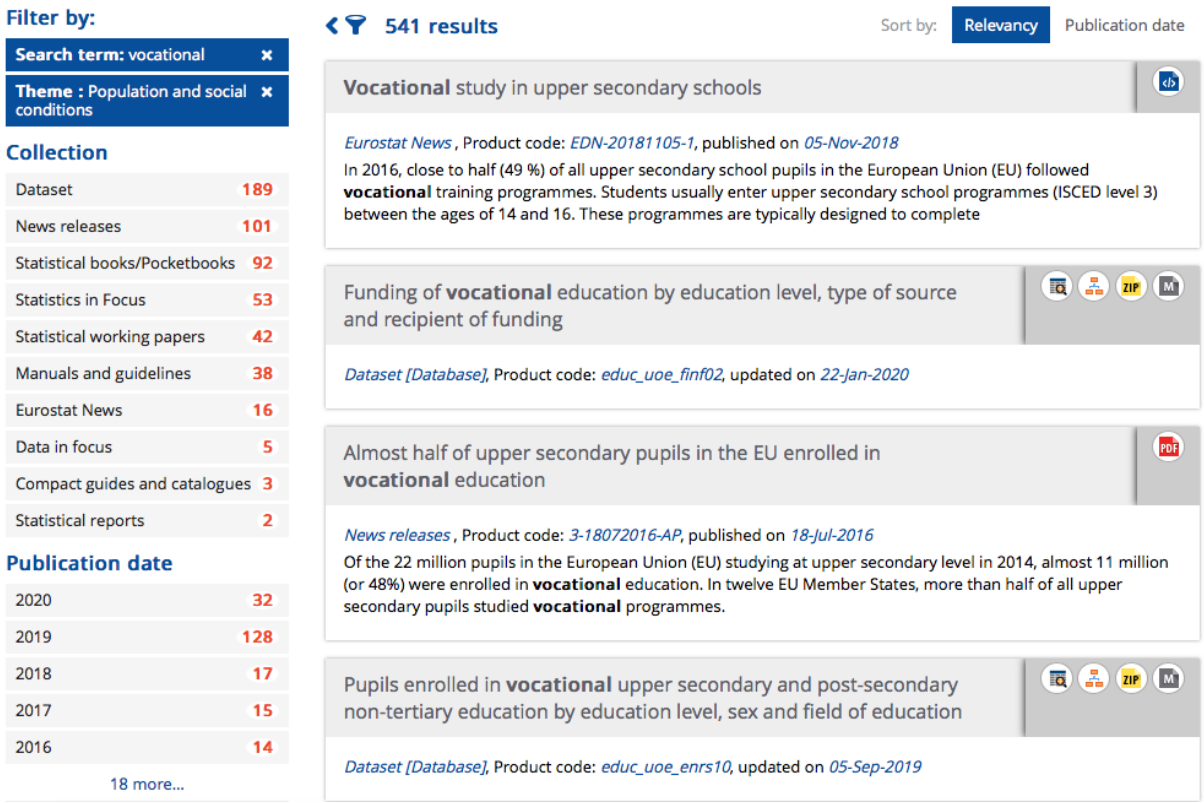

Figure 3

To narrow down the results further, the user is advised to either: a) sort by publication date (instead of relevancy) and start the search from the most recent publications or b) directly filter by publication date. When filtering by publication date, it is important to research both the year for which the tracer study is being carried out and all the subsequent years up to present, in order not to miss valuable information.

4.3. Step 3 – obtain the necessary datasets

We can filter the 541 results by publication date – for example selecting 2019, which gives 128 results (Figure 4) – and start navigating through the pages searching for interesting/relevant titles.

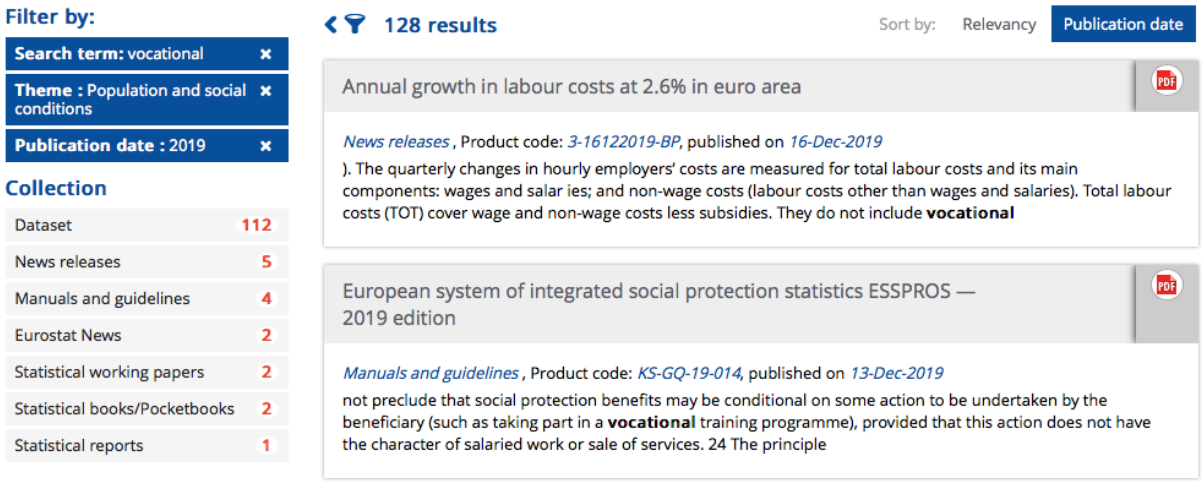

Figure 4

A very interesting dataset that appeared on the first page of our search results is the one titled "Employment rates of young people not in education and training by sex, educational attainment level and years since completion of highest level of education" (Figure 5).

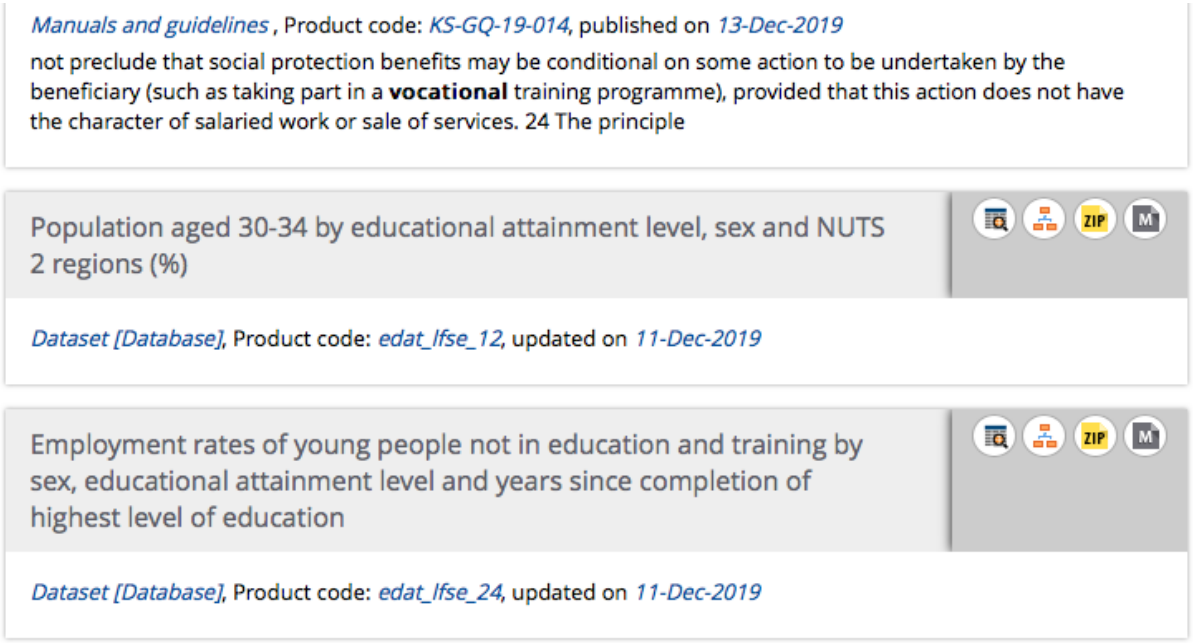

Figure 5

By clicking on the title we can open the dataset details and find information regarding its last update, the oldest data, and the most recent data it contains (Figure 6).

DATASET DETAILS

Employment rates of young people not in education and training by sex, educational attainment level and years since completion of highest level of education

View table

Download table

Show table location in data tree
Metadata

Additional information

Code: edat_lfse_24
Last update: 11/12/19
Oldest data: 2000
Most recent data: 2018
Number of values: 237865

Themes: Population and social conditions

Figure 6

The next step would be to access the data by clicking “view table”:

eurostat

Employment rates of young people not in education and training by sex, educational attainment level and years since completion of highest level of education [edat_lfse_24]

Last update: 24-02-2020

Table Customization show

TIME | GEO | Duration: From 1 to 3 years | International Standard Classification of Education (ISCED 2011): All ISCED 2011 levels | Age class: From 15 to 34 years | Sex: Total | Unit of measure: Percentage

TIME / GEO	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
European Union - 27 countrie	72.2	70.9	71.0	69.4	68.7	69.2	70.3	72.1	74.0	75.9
European Union - 28 countrie	72.4	71.4	71.3	70.1	69.7	70.4	71.5	73.0	74.7	76.4
European Union - 15 countrie	72.0	71.6	71.9	70.4	70.1	70.4	71.3	72.6	74.0	75.8
Euro area - 19 countries (fror	71.1	70.3	70.9	68.9	68.2	68.4	69.2	70.8	72.6	74.6
Belgium	75.5	74.5	73.4 (b)	73.1	72.6	72.3	72.7	77.0	76.0 (b)	76.9
Bulgaria	65.0	60.6 (b)	53.5	59.0	60.8	58.9	66.7	64.8	69.6	70.2
Czechia	82.2	78.3	76.8 (b)	78.3	77.9 (b)	78.8	80.2	84.5	87.6	88.0
Denmark	84.3	78.7	78.7	80.0	78.1	78.6	77.9	80.1 (b)	78.4 (b)	81.7
Germany (until 1990 former t	80.8	81.6 (b)	84.9 (b)	85.1	86.5	86.4	87.2	86.4	87.3	88.7
Estonia	60.1 (b)	60.8	70.2	70.0	72.7	76.3	76.7	74.1	78.8	74.5
Ireland	69.2	64.1	64.1	62.4	66.2	68.1	70.8	75.1	77.6 (b)	78.7
Greece	61.6 (b)	54.6	47.0	38.7	36.3	39.5	41.3	45.0	47.5	50.5
Spain	60.7	56.5	57.1	52.4	48.9	54.3 (b)	55.9	58.8	64.3	66.3
France	69.2	69.7	69.8	68.5	67.3 (b)	66.7 (b)	64.6	65.2	67.0	70.6
Croatia	70.1	66.9	56.8	55.1	48.0	56.1	57.6	65.6	62.6	67.3
Italy	55.5	52.0	51.9	48.8	43.9	40.6	44.2	48.4	49.7	52.0
Cyprus	78.6 (b)	75.7	70.1	68.8	59.2	65.1	66.1	70.1	67.6	76.5
Latvia	64.4	59.4	66.0	68.7	75.2	74.8	77.9	79.4	75.4	81.7
Lithuania	67.8	68.2	66.4	72.4	73.1	78.7	78.5	81.0	82.1	83.3
Luxembourg	80.9 (b)	83.5	83.3	80.1	75.2	79.2	80.1 (b)	80.2	84.6	86.0
Hungary	71.1	70.5	69.5	69.3	69.3	72.8	76.3 (b)	80.0	79.5	82.0
Malta	80.5	80.5	81.9	81.0	83.0	83.9	84.1	88.1	87.5 (b)	90.4
Netherlands	86.7	87.3	87.9	83.6	82.2 (b)	81.9	85.7	86.5	87.2	89.2
Austria	82.4	81.0	84.4	85.7	84.7	81.6	82.7	82.7	85.6	84.2

Available flags:
b break in time series c confidential d definition differs, see metadata
e estimated f forecast n not significant
p provisional r revised s Eurostat estimate
u low reliability z not applicable

Special value:
: not available

Figure 7

The table (Figure 7) shows the percentage of young people (both female and male) aged from 15 to 34 years that are employed after 1-3 years since their graduation from the highest level of education they achieved, per country and per year, and for all education levels.

The table is customisable; for example, we can select a different period of time since completion of highest education level (Figure 8).

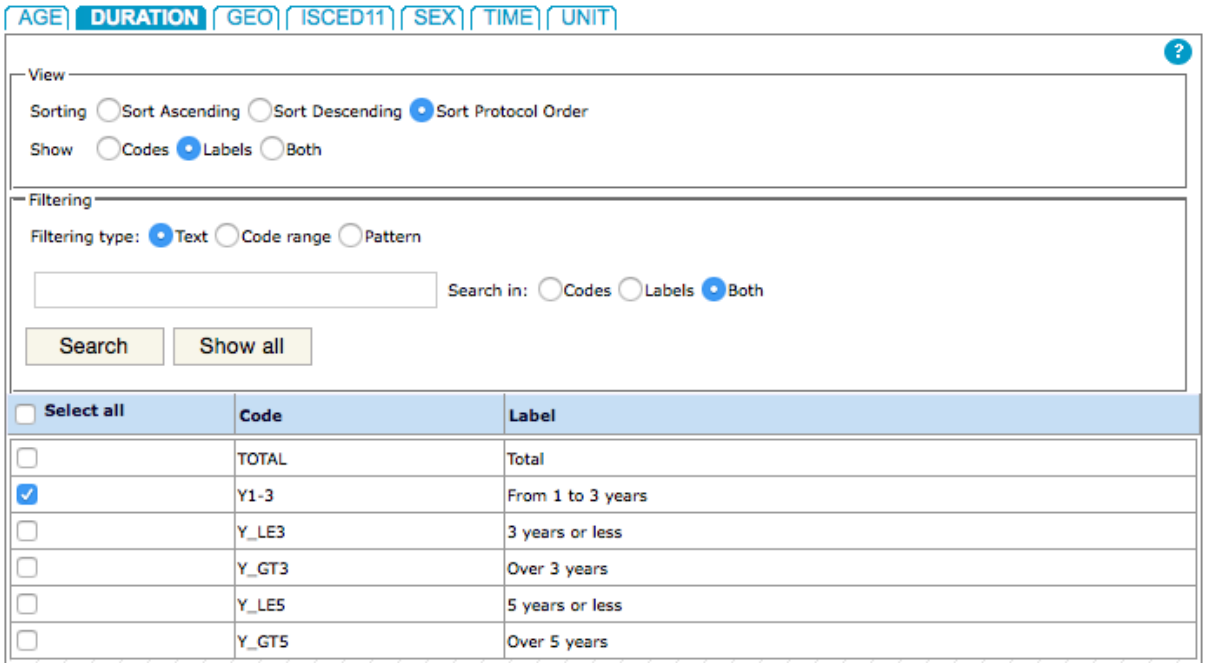

Figure 8

We can also select a different age class from the default, as well as level of education and female or male graduates. Let's assume for example that we want to find out what percentage of vocational education graduates aged between 20-34 years is employed 1-3 years after their graduation per year and per country. We need to customise the table selecting the appropriate age class and the appropriate ISCED 2011 level of education, which in the case of vocational education would be "Upper secondary and post-secondary non-tertiary education (levels 3 and 4) – vocational" (Figure 9).

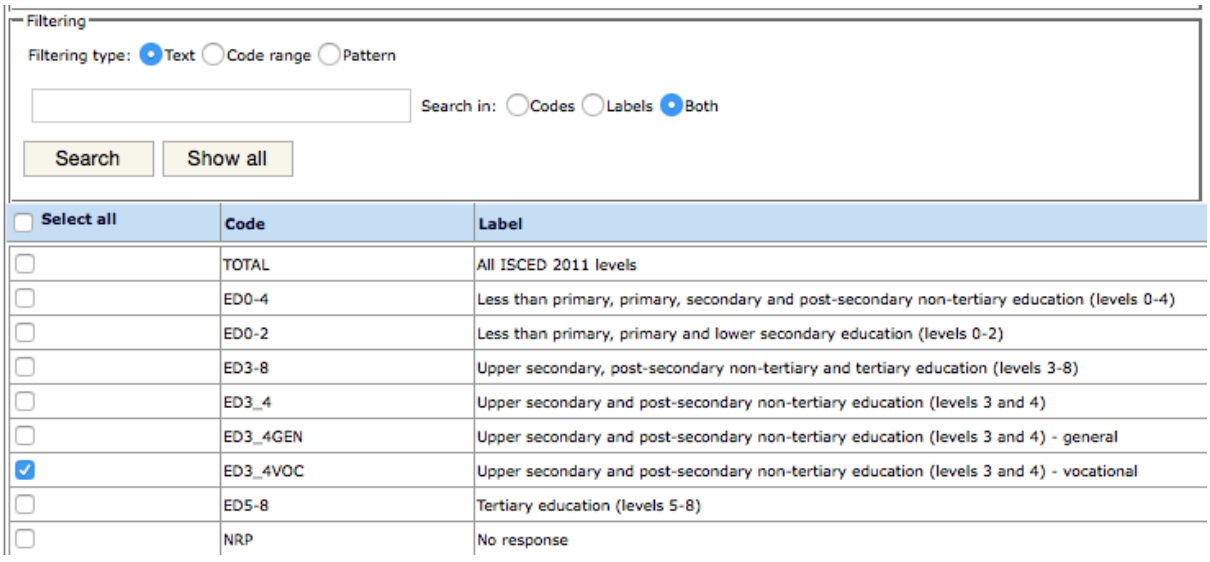

Figure 9

Once we customise the table (ISCED11: upper secondary and post-secondary non-tertiary education (levels 3 and 4) – vocational, Unit: percentage, Age: from 20 to 34 years, Sex: total, Duration: from 1 to 3 years) we can download it in any of the available formats: Excel, CSV, HTML, PC-AXIS, SPSS, TSV or PDF. To elaborate the process in this example we will use the table of the "preview option" (printable version) so as to view the data more easily (Figure 10).

TIME ▸	2014	2015	2016	2017	2018
GEO ▾					
European Union - 27	72.4(b)	72.4	75.1	76.0	79.0
European Union - 28	73.2(b)	73.0	75.3	76.6	79.5
European Union - 15	74.6(b)	74.0	75.5	76.9	79.8
Euro area - 19 countr	73.4(b)	73.0	75.1	76.0	79.0
Belgium	70.2(b)	75.6	73.1	72.2(b)	76.7
Bulgaria	63.8(b)	61.5	64.2	59.1(u)	66.4
Czechia	80.8(b)	81.5	87.1	87.5	87.7
Denmark	82.8(b)	81.6	80.7(b)	82.8(b)	85.6
Germany (until 1990	89.4(b)	90.0	90.1	91.3	92.4
Estonia	79.6(b)	83.2	82.1	86.2	76.6
Ireland	57.7(b)	61.5	72.2	71.9(b)	76.9
Greece	41.6(b)	37.5	40.7	48.4	50.5
Spain	55.3(b)	57.4	61.3	58.5	70.0
France	68.4(b)	61.5	66.7	64.0	72.2
Croatia	47.8(b)	46.7	70.3	59.4	68.8
Italy	40.2(b)	42.7	48.7	50.8	53.9
Cyprus	54.9(bu)	70.3	72.6(u)	51.8(u)	67.3
Latvia	71.5(b)	74.2	74.8	69.1	75.8
Lithuania	72.4(b)	75.6	74.9	71.5	79.2
Luxembourg	78.5(bu)	100.0(bu)	100.0(u)	84.4(u)	95.4
Hungary	76.4(b)	77.3(b)	84.4	85.9	87.1
Malta	95.4(b)	90.4	96.0	93.9(b)	91.0
Netherlands	83.4(b)	85.8	86.7	86.9	87.9
Austria	86.9(b)	84.3	85.5	89.2	87.3
Poland	67.1(b)	70.1	74.3	75.2	78.4(b)
Portugal	61.6(b)	70.1	69.8	78.9	77.4
Romania	60.0(b)	62.4	63.3	67.2	69.0
Slovenia	64.0(b)	62.4	72.3	80.8	84.5
Slovakia	70.8(b)	73.5	77.2	81.6	84.7
Finland	75.1(b)	71.6	77.1	77.0	78.5
Sweden	81.3(b)	81.6	83.9	87.8	88.0(b)
United Kingdom	79.1(b)	78.5	77.0	82.0	84.2
Iceland	89.9(b)	93.2	91.0	:(u)	:
Norway	82.2(b)	88.8	88.6	83.9	90.4
Switzerland	85.2(b)	84.2	85.6	85.1	85.7
Montenegro	40.1(b)	48.9	45.6	53.0	53.6
North Macedonia	:	:	42.1	47.9	45.4
Serbia	39.4(b)	44.4	51.4	58.2	59.3
Turkey	53.1(b)	59.3	55.1	55.5	54.8

Figure 10

The above table contains very useful information regarding the effectiveness of vocational education in terms of employability in each European country. The data about graduates that we obtain from Eurostat at EU-level can be compared with information from other sources, such as graduate surveys

at national/local level, as well as facts about the labour market, industry, and economy. The combination and comparison of data from various sources can help VET providers connect the results with aspects of the quality of vocational education.

For example, the table shows that the percentage of VET graduates that were employed 1-3 years after their graduation in Greece in 2015 was 37.5%, which is significantly lower than the percentage in most European countries. This can be partially explained by the overall situation of the economy in 2015, when Greece was still suffering from the financial crisis and its impact on the labour market. The percentage started rising from 2016 onwards reaching 50.5% in 2018, so the gradual improvement of the economy probably had a positive impact on the employability of VET graduates. We can repeat the process filtering by another level of education, for example tertiary, to see if the percentages were similar for all types of graduates and confirm our conclusions. In case the percentages for different educational levels have significant deviations (e.g. if the percentage of employed VET graduates is more than 10 percentage points lower than the respective percentage of tertiary education graduates), this could be related to the quality of VET provision and the relevance of curricula with the requirements of the labour market. The same applies if for example the percentage in some EU countries is significantly lower compared to other EU countries.

5. Sorting and storing of tracking data

It is recommended to VET providers to create an internal system of sorting and storing the data that are necessary for graduate tracking studies (or other purposes such as promotional activities), so as to not overload the institution's database with an unnecessarily large amount of statistical datasets and to be able to find relevant information easily.

Create an internal database to store the data

The provider is advised to start by developing a database where the necessary tracking information will be saved. This database can be as simple as a table (e.g. an Excel sheet) with specific pre-defined variables that will be updated per year with the statistical data that have been extracted from the open public datasets.

The provider should also define the variables that will be monitored, according to which information would be more useful and relevant to the benchmarking scope of this tool, i.e. comparing outcomes of the institution's own graduate tracking survey against national and European statistics. It is recommended to include the following minimum variables:

- Employment rate of vocational education graduates (e.g. 1-3 months after graduation, 6 months after graduation, 1 year after graduation etc)
- Employment rate of general education graduates
- Rate of female vocational education graduates
- Rate of male vocational education graduates.

Keep the database up-to-date

The variables of the database should be updated at least once per year, by following the steps of the process of chapter 4 "Extracting comparable data about VET graduates".