View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Warwick Research Archives Portal Repository

Manuscript version: Published Version
The version presented in WRAP is the published version (Version of Record).

Persistent WRAP URL:
http://wrap.warwick.ac.uk/148593

How to cite:
The repository item page linked to above, will contain details on accessing citation guidance from the publisher.

Copyright and reuse:
The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions.

Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRAP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Publisher's statement:
Please refer to the repository item page, publisher's statement section, for further information.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk

warwick.ac.uk/lib-publications

Department for Education

Enhancing a labour market information database: LMI for All

Stakeholder Engagement and Usage, Data and Technical Developments (2019-2020)

Research report

February 2021

Sally-Anne Barnes, Jenny Bimrose, Jeisson Cárdenas-Rubio, Rob Wilson, David Owen, Terence Hogarth, Luke Bosworth and Rosie Day – Institute for Employment Research, University of Warwick

Graham Attwell and Philipp Rustemeier – Pontydysgu

Contents

1. Introduction 1

1.1. The development of LMI for All: An overview 2

1.2. Programme of work, 2019-2020 2

1.3. Report structure 3

2. Technical infrastructure and developments 4

2.1. Objectives and outcomes 4

2.2. LMI for All Infrastructure 5

2.3. Server capacity and API usage 5

2.4. IT infrastructure developments 7

2.5. Renewal of the LMI for All website 7

2.6. Development of a new widget: Skillsometer 8

2.7. Summary and next steps 11

3. Data developments 12

3.1. Objectives and outcomes 12

3.2. Overview of the LMI for All database, organising framework and approach to providing data 12

3.3. New and updated data 17

3.3.1. New mappings 18

3.3.2. Course data 19

3.3.3. Vacancy dataset 20

3.4. Implications of the change from SOC2010 to SOC2020 25

3.5. Summary and next steps 26

4. Stakeholder engagement 27

4.1. Objectives and outcomes 27

4.2. Stakeholder plan 28

4.3. Stakeholder activities and dissemination of the LMI for All service 29

4.4. Engagement through the LMI for All website and social media 31

4.5. Supporting careers practitioners 32

4.6. Take-up of LMI for All and Careerometer 34

4.7. Summary and next steps 34

5. Future developments and next steps 36
References 39
Annex A: LMI for All reports 41

List of figures

Figure 2.1 Long-traffic to LMI for All, June – September 2020 6
Figure 2.2a Screenshot of Skillsometer welcome card (start of quiz) 10
Figure 2.2b Screenshot of Skillsometer quiz (selected example) 10
Figure 2.2c Screenshot of results from Skillsometer quiz – occupational description 10
Figure 2.2d Screenshot of results from Skillsometer quiz – pay and hours data 11
Figure 3.1 Percentage of queries by dataset, October 2019-September 2020 14
Figure 3.2 Overview of data update cycle and new data coming online 18
Figure 3.3 Steps for integrating the job portals information with the LMI for All infrastructure 21

List of tables

Table 3.1 LMI for All Data overview, September 2020 15
Table 3.1 continued, LMI for All Data overview 16
Table 3.2 Structure of vacancy dataset under construction as of September 2020 24
Table A.1 Overview of LMI for All development 41

1. Introduction

Labour market information (LMI) has been long recognised as a key resource throughout the careers stakeholder community in the UK (Alexander, McCabe, and De Backer, 2019; Barnes, 2015; The Careers & Enterprise Company and WorldSkills UK, 2016) and internationally (CEDEFOP, 2016; Mezzanzanica and Mercorio, 2019; OECD, 2015). In response to the need for a consistently robust and reliable source of LMI to support career transitions, the LMI for All service was first piloted and developed from 2012 in the UK with funding from the UK Commission for Employment and Skills (UKCES), and subsequently by the Department of Education from 2016. The overall aim of LMI for All is to provide a single access point for multiple, high quality sources of national and regional data that is openly accessible with the primary purpose of supporting individual labour market transitions and career decision making with the provision of relevant and robust LMI. Over time, the service has responded to a number of policy issues and concerns. For example, the Careers Strategy identified 'LMI for All' as a reliable source of LMI to support career transitions and enable access to information on opportunities (Department for Education, 2017). This followed on from a recommendation from Gatsby that 'LMI should be available to all schools through the 'LMI for All' service (The Gatsby Foundation, 2014, p.11).

A profound change in policy priorities, however, came with the onset of the COVID-19 pandemic in early 2020. This has transformed labour markets around the world. In parallel with these seismic international shifts, the context in which LMI for All is being developed and delivered in the UK has changed dramatically, with a policy focus on the emerging and predicted high levels of unemployment, with an urgent need to support entry and re-entry into employment. Careers advice is part of the re-framing of policy. On 8th July, the Chancellor of the Exchequer delivered a budget speech in the House of Commons, entitled 'A plan for jobs', in which he stated:

> …the evidence says careers advice works, so we will fund it, with enough new careers advisers to support over a quarter of a million more people (Sunak, 2020).

In this same speech, a number of policy initiatives were announced by the Chancellor (such as the Jobs Retention Scheme and the Plan for Jobs), many designed to support people into, or back to, paid employment. Successful implementation of these policy initiatives will require continued and easy access for career practitioners and their clients/customers to high quality, up-to-date LMI that supports individuals effectively in their transitions into and through suitable job and training opportunities.

The LMI for All service is widely used by a range of beneficiaries and the enhancements to the service and database over the last year ensures that it continues to be a key place to access up-to-date and robust data on the UK national and regional labour markets.

Current and future enhancements to LMI for All will extend the service to enable access to some local level data.

1.1. The development of LMI for All: An overview

A consortium led by the Institute for Employment Research at the University of Warwick has operated the LMI for All service with Pontydysgu, an educational technology organisation, since 2012.[1] A feasibility study in 2012 proved the viability of a labour market information database; it was then refined and further developed up to 2015. This included improvements to the technical infrastructure and the addition and expansion of datasets (see section 3.2 for details of the datasets). In 2015, the LMI for All service was given full project status and since then the project has continued to be developed with technical improvements, enhancements to the datasets, and a variety of stakeholder engagement activities. The aim of these stakeholder activities has been to increase awareness of the service, support developers in the utilisation of the database, and support career practitioners' knowledge of labour market information and its use in practice. An overview of the development phases of LMI for All and associated reports are presented in Annex A.

A new phase of work to develop and enhance the LMI for All service was commissioned by the Department for Education in 2018. This annual report is the second in a series of three reports documenting the programme of work being undertaken over a three year period (2018 – 2021). The overall aim of LMI for All continues to be to increase and widen the use of high-quality labour market information in order to support decisions about careers and learning.

1.2. Programme of work, 2019-2020

This report details work undertaken to support the delivery and enhancement of the LMI for All service from October 2019 to September 2020. As in previous years, research and development over this period has comprised three interrelated work streams: (i) technical; (ii) data; and (iii) stakeholder engagement. The objectives for the three-year programme of work are:

- To maintain a comprehensive, high quality data offering that can inform career choices;

[1] The consortium is referred to as 'the team' throughout this report.

- To maintain the supply of data to current and future third-party users thereby meeting the requirement to provide open access to government data as set out in the eight Principles of Open Data[2];
- To offer a technical infrastructure that provides a secure, engaging, accessible and reliable platform for LMI for All;
- To promote the widest possible take-up among third-party websites and applications as a means of opening-up the data to individual decision-makers; and
- To build wider awareness, understanding and support for LMI for All among key stakeholder groups, including policy, careers and technical communities.

1.3. Report structure

Following the structure of the first report in this series, this report comprises three main sections following the three interrelated work streams, namely technical, data and stakeholder engagement. The objectives for each work stream are set out alongside the outcomes achieved in this reporting year, October 2019 to September 2020.

Section two includes an overview of the LMI for All technical infrastructure, presents enhancements to the infrastructure and provides a review and assessment of the infrastructure in terms of server capacity and usage. The development of online learning units and a new widget, Skillsometer, now part of the service are also detailed in this section.

Section three on data development includes an overview of: the structure of the LMI for All database; the methodology used to provide data to the database; and the datasets and indicators within the database. This section also includes information on the enhancements to the database and the current work around utilising web scraping and big data analysis techniques to extend the database.

Section four includes details on stakeholder engagement activities undertaken over the last year to promote the LMI for All service. It reports on the various events, the shift in the mode of delivery in response to COVID-19, as well as take-up of the service.

Section five sets out the future programme of work for the LMI for All service in 2020-2021.

[2] The Eight Principles of Open Data, include: complete; primary; timely; accessible; machine processable; non-discriminatory; non-proprietary; and license-free (for more information, see: https://opengovdata.org/)

2. Technical infrastructure and developments

This section of the report provides an overview of the current IT infrastructure of LMI for All and outlines enhancements to the technical infrastructure and developments. It also includes details on server capacity and usage over the last year. Additionally, this section provides information on updates to the LMI for All website. Finally, the development of a second widget, Skillsometer, which is now provided as part of the service is detailed in this section.

2.1. Objectives and outcomes

For the three year programme of work, there are four key objectives guiding the provision and enhancement of the technical infrastructure, plus support for the LMI for All service. The outcomes achieved in this reporting year are set alongside each objective.

- To maintain a secure and robust infrastructure for the data, including test and production servers and provision for backup services. For this, the hardware and software environment has been updated, maintained and monitored ensuring a future-proof infrastructure. The infrastructure has coped well with spikes in usage and increases in the size of the database. Whilst the infrastructure can be extended, this has not been necessary over the last year.
- To implement processes to ensure that the data made available to developers and end-users meets required quality standards. Data continues to be made available through the LMI for All API,[3] which is freely available. Quality procedures are in place to check data before it is loaded to the database and published, and the APIs have been programmed to ensure that data providers' rules are in place. These rules mean that data are suppressed if disclosive or unreliable due to sample size.
- To provide modern and flexible software tools to allow the querying of the database by external users. Tools that are in place to enable external users or developers to query the database continue to be current and flexible.
- To provide software tools and spaces for documenting the process and allowing public access to the LMI for All database. The website, which was redeveloped in 2018-2019, has been further enhanced and refined to support developers and practitioners looking for LMI.

[3] An API or application programming interface is a computing interface that enables access to and defines requests to the database. It has been custom built for LMI for All. An API is a set of routines, communication protocols, and tools for building software applications. API queries produce formatted responses into a universal data format known as JSON (JavaScript Object Notation), before being sent back to the developer/client.

2.2. LMI for All Infrastructure

The development of the LMI for All infrastructure is documented in earlier reports (see Annex A). As demand for the service has increased over the last few years, the infrastructure was improved to not only increase capacity, but also to manage any future demands. In short, the LMI for All infrastructure uses a UK-based secure cloud service. The infrastructure comprises:

- A data server (production environment) which stores the data, processes queries from the API and serves data;
- A development data server (test environment) which enables new functionality to be tested as well as checking the integrity of new or updated data; and
- A monitoring system to ensure the system is running efficiently.

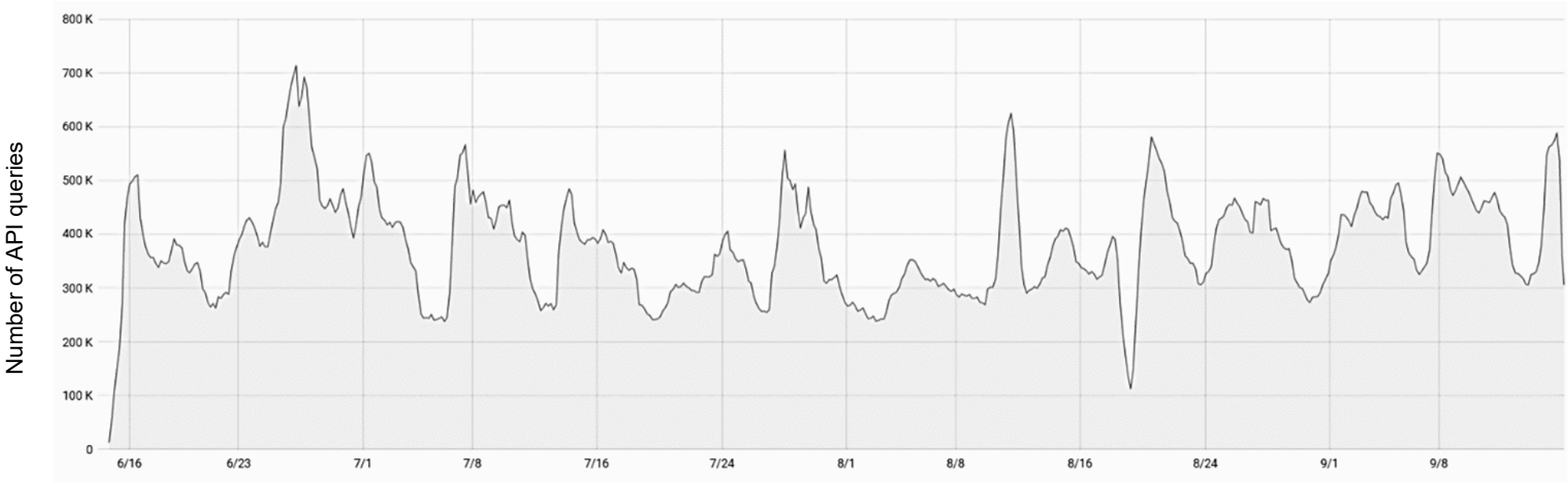
Pivotal to the infrastructure are the two API servers and load balancer. The APIs provide external users access to the database and have rules in place to filter and restrict access to disclosive and unreliable data. As part of the quality procedures, these rules are reviewed when data are updated to ensure they are both current and relevant. The system is designed to automatically deploy additional API servers if and when required to ensure continuity and consistency in the service. The load balancer provides an additional layer of security to our servers and optimises response query time by managing and distributing queries to the API servers.

2.3. Server capacity and API usage

Over the last year, there has been a significant increase in the use of the LMI for All service. Usage data shows that there continues to be variations in terms of time of day and whether it is a weekday or weekend; this is to be expected. Figure 2.1, below, shows the number of queries to the LMI for All service over a six month period.

Current data on usage suggests that there are on average 500,000-600,000 queries per day (in 2018-2019, this was on average 400,000-500,000 queries per day), which equates to 7-8 queries a second (see Figure 2.1). The LMI for All API monitoring system reports regular spikes of around 700,000 queries a day, which the system is able to manage.

Figure 2.1 Long-traffic to LMI for All, June – September 2020



Source: LMI for All API Monitoring system, 14/06/20-14/09/20

2.4. IT infrastructure developments

At present, the IT infrastructure is managing current queries and short-term spikes in usage so there are no plans to make any changes.

During the last year, there has been an estimated 99.5 per cent uptime for the API and two instances of failure in the broader system which interrupted service for some LMI for All users. As in the previous reporting year, a certification expiration resulted in interrupted service for those connecting using a certificate for a 24 hour period before the problem was identified. The monitoring system implemented last year did not report the error promptly. Therefore, a new monitoring system has been purchased and implemented to address certificate errors. The second failure was due to a power outage at the server farm.

In addition, and in response to this interruption, a technical issues contact form has been embedded in the front page of the LMI for All website. Any issues or queries reported in this form are emailed to three members of the LMI for All team to ensure issues are identified and resolved promptly. User queries and problems continue to be addressed through this and the main email address.

The extract transform and load (ETL) system has been developed for the pay and hours dataset and work is well underway for additional datasets in LMI for All. The ETL will help automate the data updating process ensuring that data have fixed file format, structures and variable names. This will form part of the data quality checks undertaken when datasets are refreshed.

2.5. Renewal of the LMI for All website

In 2018-2019, the LMI for All website structure was significantly revised with a new template and design that continues to be used across all dissemination materials. The website structure was revised to ensure the service being offered was clear in directing users to the section of the website appropriate to their query. In addition, the template needed updating so that the team could take advantage of new plug-ins, such as the newsletter plug-in. Work on refining and updating the content has continued this year, as well as the addition of new content including some short online learning units on LMI. Behind the scenes the LMI for All web server has been updated using an advanced operating system (NixOS), which is faster, more secure and will enable the LMI for All team to use the new plug-ins. The server will automatically update to the latest versions of PHP and MySQL.

2.6. Development of a new widget: Skillsometer

As part of the LMI for All service, Careerometer (a free widget) was developed a few years ago. This was developed and made freely available, supporting those unable to develop their own interface to LMI for All as they may lack the technical capacity and/or resources. The Careerometer widget code can be copied by users and embedded in their own websites. It has been designed specifically for young people and enables access to key occupational data. There has been considerable take-up of the Careerometer widget since its launch, as over 160 institutions have implemented it in their website. Careerometer has been successful as it provides access to simple data by occupation and enables users to compare occupational information and data. Careerometer has been maintained over the last year and promoted as one way in which to access LMI for All data. Whilst, Careerometer was developed specifically for schools, it has proven popular with a range of careers and learning providers. Due to the success of Careerometer, it was proposed that a second widget be developed similarly aimed at schools and colleges; Skillsometer was developed and released in August 2020. It is a quiz-based widget that once users complete, they are presented with a number of occupations matching their responses.

Suggestions for this second widget were gathered during stakeholder engagement activities and through social media with a range of ideas generated. Initial ideas for a widget with more of a regional focus were explored. Whilst regional data were available and it was possible to design a pathway through the data, the structure and format of a widget restricted the possibilities of a regional based widget. Widgets have dynamic JavaScript so space and functionality can be restricted. A benefit of a widget is that it can be easily embedded within a platform, which means it is accessible for those with limited technical skills. For Skillsometer, similar to Careerometer, all that is needed to include it on a web page is to add one line of html code.

Drawing upon feedback and in consultation with the Department for Education, it was agreed that a widget drawing upon the US O*NET data would be useful. The O*NET dataset (mapped to UK Standard Occupation Classification as part of the LMI for All project) links abilities, skills, interests and knowledge to each occupation. It is an extensive dataset that is underused within LMI for All due to the complexity of visualising the information and understanding the concepts used. A number of applications and websites that use the LMI for All service tend to start with the idea that users are aware of different occupations and search for information by occupation. However, for those unsure of what jobs (or occupations) they may be interested in, a new starting point is needed. The O*NET data provides this opportunity through ranking or rating different interests, which are associated with different occupations.

The front-end of the widget was developed with young people in mind. This design was based on cards with emoji as a form of rating the statement on each card. This initial design was tested with five teenagers. Feedback was generally positive with some helpful suggestions regarding the instruction card and how cards could be designed. Whilst there had been concern around the use of emoji, young people liked this visual representation and understood what each emoji meant. The widget was developed and a further pilot was undertaken with a larger user testing group from the National Careers Service. Feedback was again positive with suggestions for refining the presentation of information and how it might be used alongside Careerometer. Final revisions were made and it was agreed that the widget would be called Skillsometer (see Figure 2.2a). Screenshots of the final version are shown in Figure 2.2a-d.

The following describes Skillsometer and is suggested text for those wishing to embed and introduce Skillsometer on their website:

> Skillsometer can help you discover what jobs you might like to do in the future. You will be presented with a series of statements. Select the emoji that shows how you feel about each statement. You will then be given suggestions of jobs linked to what you most enjoy doing. For each job, you will be presented with some headline UK data.

Skillsometer has been designed for those who are not sure what jobs they may be interested in. Thinking about skills, interests, and the ways these can link to jobs is recognised as a helpful first step in identifying possible future jobs. Those that take the quiz are required to reflect on a number statements and decide what they love, are not sure about or dislike (see Figure 2.2b). The statements are presented within six well established occupational categories (Artistic, Realistic, Investigative, Conventional, Enterprising, and Social) (see Holland, 1997), which are then ranked against jobs. Once the user has completed the quiz, they are given a short list of job suggestions that are most likely to be suited to their own particular skills and interests (see Figure 2.2c). Each job can be explored as a description is presented together with information on pay and hours (see Figure 2.2d). For further information on the job, and to compare with other jobs, the user is then encouraged to use Careerometer.

Figure 2.2a Screenshot of Skillsometer welcome card (start of quiz)

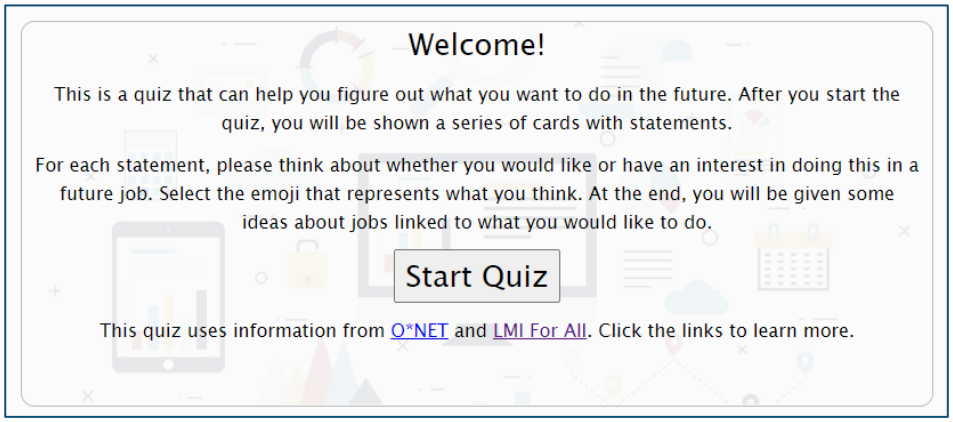


Figure 2.2b Screenshot of Skillsometer quiz (selected example)

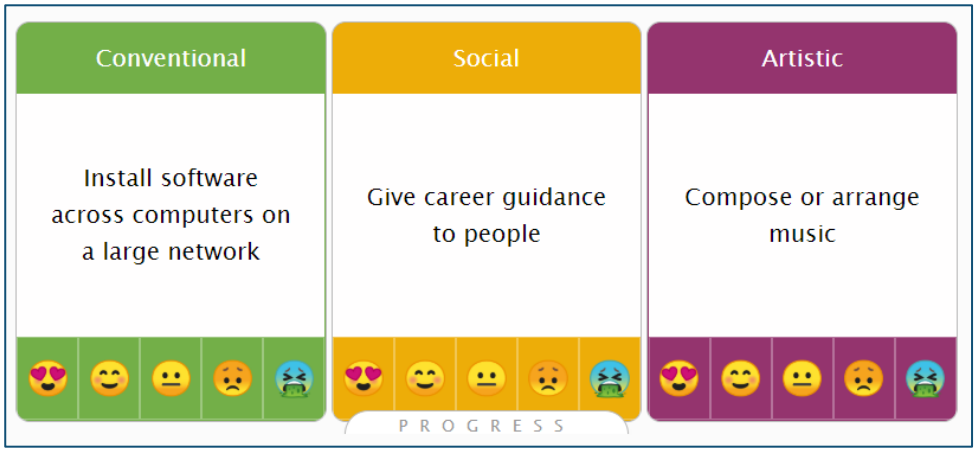


Figure 2.2c Screenshot of results from Skillsometer quiz – occupational description

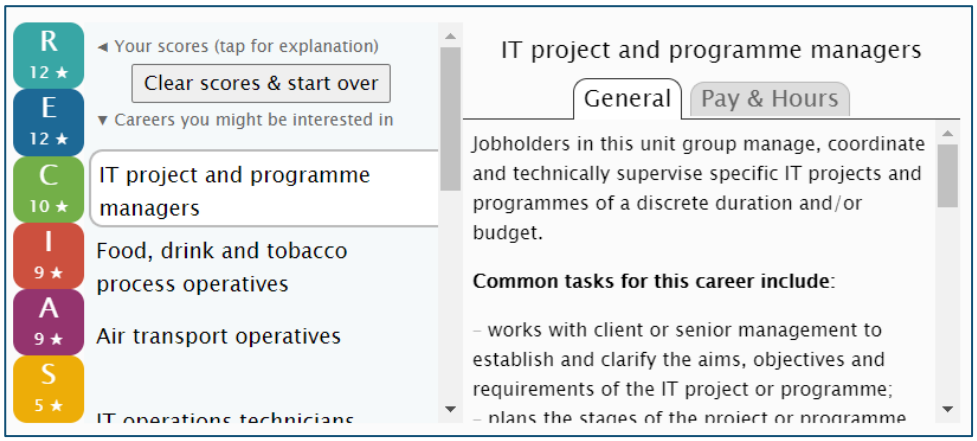
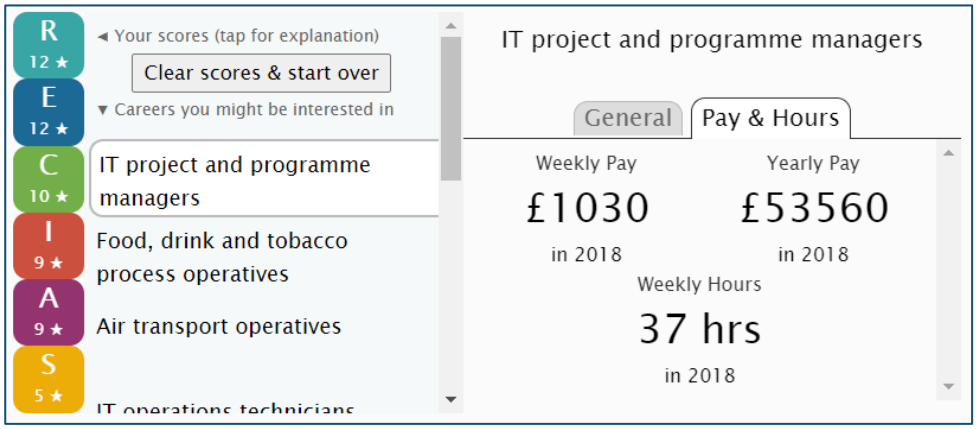


Figure 2.2d Screenshot of results from Skillsometer quiz – pay and hours data



Skillsometer was released in August 2020 and is a freely available widget. It will be promoted through stakeholder engagement activities and social media.

2.7. Summary and next steps

During this reporting period, the IT infrastructure has been maintained and monitored. As usage of the service has increased, the infrastructure has been reviewed to ensure it remains current and efficient. Improvements to the website and web server have enhanced the outward facing part of the service. The major part of the technical work has been the development of Skillsometer, plus preparing the database and API ready for new datasets which are detailed in section 3.

Over the next year, work will continue on developing the ETL system, close monitoring of the IT infrastructure, and the new learning units will be further developed as they offer useful support for those working with LMI. Significantly, the new datasets to be added to the database will need to be closely monitored in terms of operation and uptake.

3. Data developments

This section provides details of the data and data refreshing undertaken for the LMI for All service during this reporting year. It includes a brief overview of the database, organising framework and methodology. Details on the data refreshed during the reporting period are provided, together with information on the new datasets being added to the database.

3.1. Objectives and outcomes

This strand of work is about the management of the data and datasets available in the LMI for All database. This includes refreshing and updating the data, monitoring the rules for each dataset, undertaking quality checks and managing data documentation. As in the previous reporting year, the objectives of the data development strand have focused on the:

- Refreshing existing datasets held in LMI for All. Following a schedule of updating, data have been refreshed and new datasets added.
- Management of documentation relating to the datasets. Information on each dataset are provided on the LMI for All website in a dedicated section. This is reviewed and refreshed when new data are available to ensure they are current.
- Quality assurance and checking, including checks to ensure data rendered by the API are consistent with values contained in the dataset generated. Quality procedures were implemented during the early stages of the project. These are followed by the data and technical teams and regularly reviewed.

3.2. Overview of the LMI for All database, organising framework and approach to providing data

An overview of the structure of the database, organising framework and the methodology and approach to providing data is well documented in earlier reports (see Annex A).

The LMI for All database uses occupations classified at the unit group 4-digit UK SOC2010 as its organising framework.[4] This includes the following dimensions and characteristics:

[4] For more information on the UK Standard Occupational Classification (SOC), see: https://www.ons.gov.uk/methodology/classificationsandstandards/standardoccupationalclassificationsoc/soc2010

- 369 detailed occupational categories (SOC 2010 4-digit level);
- 75 detailed industries (roughly equivalent to SIC 2007 2 digit level);
- Employment status (full-time, and part-time employees and self-employment);
- Highest qualification held (equivalent to the 9 levels of the National Qualification Framework (NQF)[5]);
- Countries within the UK and English regions; and
- Gender and age.

LMI for All is built upon an index of around 28,000 job titles mapped to SOC2010. Contextual information about occupations are included in the database from two sources: occupational descriptions of 4-digit occupations (based on the Office for National Statistics (ONS) information); and the skills, abilities, knowledge and interests data mapped to 4-digit occupations based on the US O*NET information and mapped to UK SOC. The LMI for All service provides access to the following datasets:

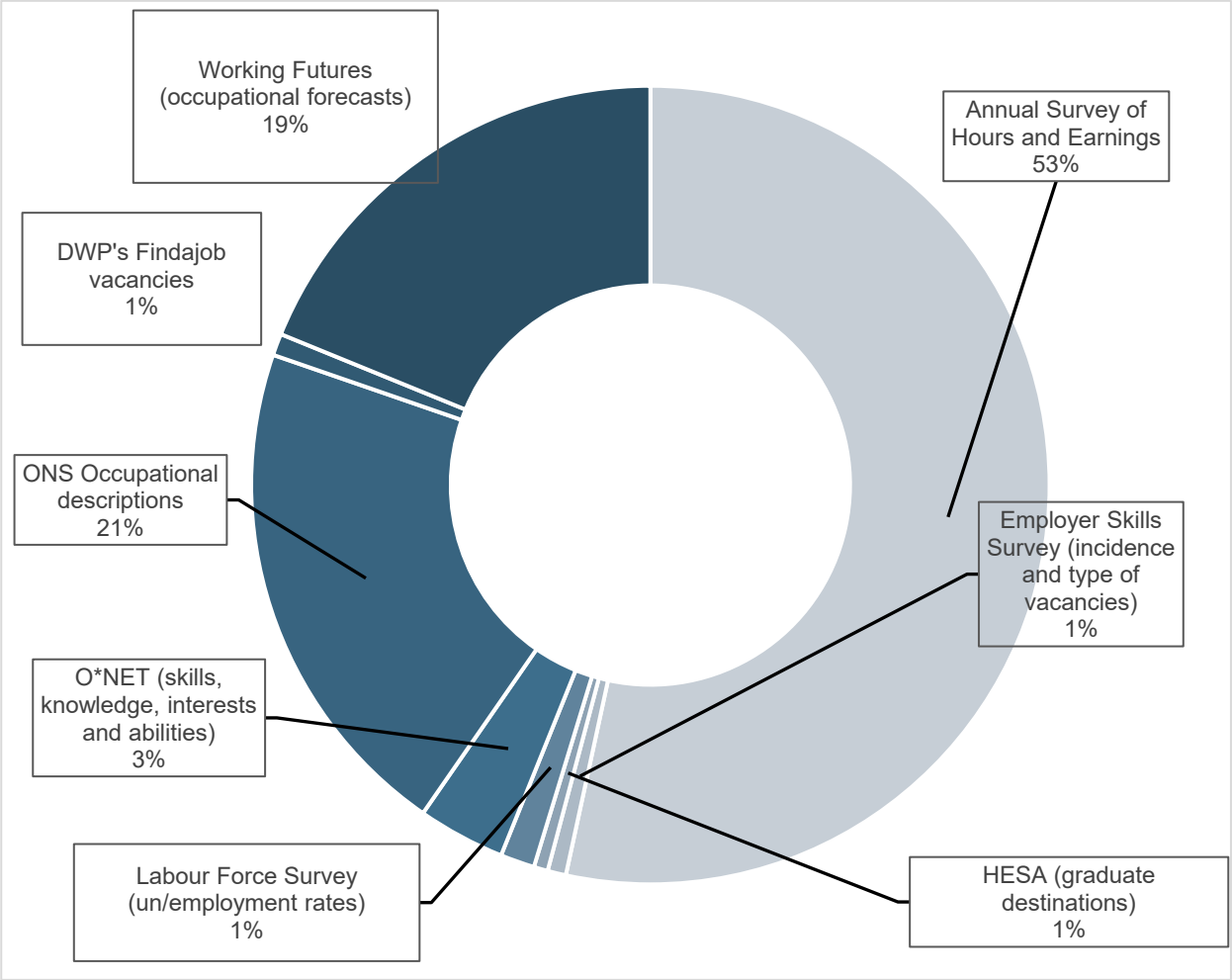
- Actual and projected employment trends from 1990 to 2027 (from the Working Futures dataset);
- Projected number of workers needed by employers to replace those retiring between 2014 and 2027 ('replacement demand') (from the Working Futures dataset);
- Pay (for employees only) (from the Annual Survey of Hours and Earnings (ASHE) dataset);
- Hours worked (for employees only, not the self-employed) (from the ASHE dataset);
- Changes in pay by detailed 4-digit occupation (modelled on ASHE and Labour Force Survey (LFS) datasets);
- Unemployment rates (based on the LFS dataset);
- Vacancy data by occupation and industry (from the Employer Skills Survey (ESS) dataset); and
- Geographical patterns of employment and travel to work distances (from the 2011 Census of Population).

[5] For more information on the NQF, see: https://www.gov.uk/what-different-qualification-levels-mean/list-of-qualification-levels

Job vacancies posted on the Department for Work and Pension's Findajob website is pulled through into LMI for All using the Findajob API. Query results are returned using a fuzzy matching approach in the API.

Figure 3.1, next, uses data from the LMI for All API monitoring system to report on the usage of each dataset in terms of number queries for this reporting year. As in previous years, the Annual Survey of Hours and Earning dataset (53% of all queries) continues to be a popular dataset, as well as the ONS occupational descriptions (21%) and the Working Futures occupational forecast dataset (19%).

Figure 3.1 Percentage of queries by dataset, October 2019-September 2020



Source: LMI for All API Monitoring system, based on queries made between 01/10/19-14/09/20

Table 3.1 provides an overview of the data and indicators available in the LMI for All database taking into account when data will be refreshed during this reporting period.

It should be noted that the methodology and approach to providing data has remained relatively unchanged over the last reporting year. However, new datasets have become available, mappings developed and new web scraping techniques and big data analysis are being used to construct new datasets. Over the next reporting year, these new datasets will be released to enhance the service. It will also be necessary to begin to address the challenges posed by the move to classify labour market data using SOC2020 rather than SOC2010 (see section 3.4 for more information). Table 3.1 LMI for All Data overview, September 2020

Table 3.1 LMI for All Data overview, September 2020

Dataset	Source	Indicator	Dimensions*	Period	Available updates (if known)
Employment (Historical)	*Working Futures*/ Business Register and Employment Survey (BRES)/ Labour Force Survey (LFS)	Number of jobs (employee, self-employed)	Occupation, Industry, Qualification, Geography, Gender, Status	2002-2019 #	every 2-3 years
Employment (Projected)	*Working Futures*/ Business Register and Employment Survey (BRES)/ Labour Force Survey (LFS)	Number of jobs (employee, self-employed)	Occupation, Industry, Qualification, Geography, Gender, Status	2020-2027	every 2-3 years
Employment (Replacement Demand)	*Working Futures*/ Business Register and Employment Survey (BRES)/ Labour Force Survey (LFS)	Number of jobs openings between selected years (employee, self-employed)	Occupation, Industry, Qualification, Geography, Gender, Status	2020-2027	every 2-3 years
Pay and earnings	Annual Survey of Hours and Earnings (ASHE)/ Labour Force Survey (LFS)	Average full-time earnings; plus indicative estimates of earnings distribution medians and deciles	Occupation, Industry, Qualification, Geography, Gender, Status, Age	2013, 2015, 2017, 2018 ##	annually

Table 3.1 continued, LMI for All Data overview

Dataset	Source	Indicator	Dimensions*	Period	Available updates (if known)
Hours	Annual Survey of Hours and Earnings (ASHE)	Average weekly hours	Occupation, Industry, Geography, Gender, Status	2015, 2017, 2018 ##	annually
Unemployment Rates	Labour Force Survey (LFS)	ILO Unemployment rate	Occupation, Industry, Qualification, Geography, Gender, Status	2011-2019	annually
Number of vacancies	Employer Skills Survey (ESS)	Number of vacancies, Hard-to-fill vacancies, Skills shortage vacancies, Occupation	Occupation, Industry, Geography	2011, 2013, 2015, 2017, 2019	every 2-3 years
Skills, Knowledge, Abilities, Interests	O*NET database (version 25)	Skills, Knowledge, Abilities, Interests	Occupation	2020	every 2-3 years
Occupational descriptions	ONS Standard Occupational Classifications	Structure and descriptions of occupations	Occupation	2010 (new and revised job titles amended in 2019)	only required when SOC is updated
Current vacancies	Findajob (DWP)	(Available through fuzzy search, SOC mapping)	Occupation	current	constant
UK Census Population	UK Census Population (England and Wales)	Information on geographical patterns of employment and travel to work distances	Occupation (1,2,3 digit), Geography**	2011	every 10 years
Higher education destinations	HESA	Destination of graduates immediately after graduation	Occupation, Qualification, Qualification required for job, Subject of study	2012/13-2015/16	annually
Apprenticeships	Apprenticeship Service	Apprenticeship data and frameworks	Occupation	(currently being tested)	constant

Table 3.1 Notes:
* Occupation (SOC2010 4-digit), Industry (SIC2007, 75 industries), Qualification (NQF 0-8), Geography (UK countries and English regions), Gender, Status (full-time or part-time employee and self-employed).
** Geography available for Output Areas, Lower and Middle Super Output Areas and the hierarchy of local government areas from wards to regions and nations
# For 2000-2013 data are only available at the SOC2010 2-digit level
Estimates of Pay are also available for 2013 and 2014. However, comparisons of detailed estimates between years will not provide robust information at the 4 digit level. A more limited "change indicator" is available for those interested in changes over time.

3.3. New and updated data

The following datasets have been refreshed during the reporting year:

- New employment and forecast data (2017-2027) (Working Futures);
- Unemployment rates (LFS);
- Occupational structure and descriptions (Office for National Statistics); and
- Skills, knowledge, abilities and interests (O*NET)

The database has been enhanced by two new mappings: an O*NET to UK SOC2010, and UK SOC2010 to O*NET mapping; and apprenticeship standards to multiple UK SOC2010 codes where appropriate.

The following new datasets will be made available in the next reporting year:

- Education and Skills Funding Agency course database;
- Apprenticeship vacancies; and
- A pilot dataset of UK job vacancies.

Each of these are explained in more detail below.

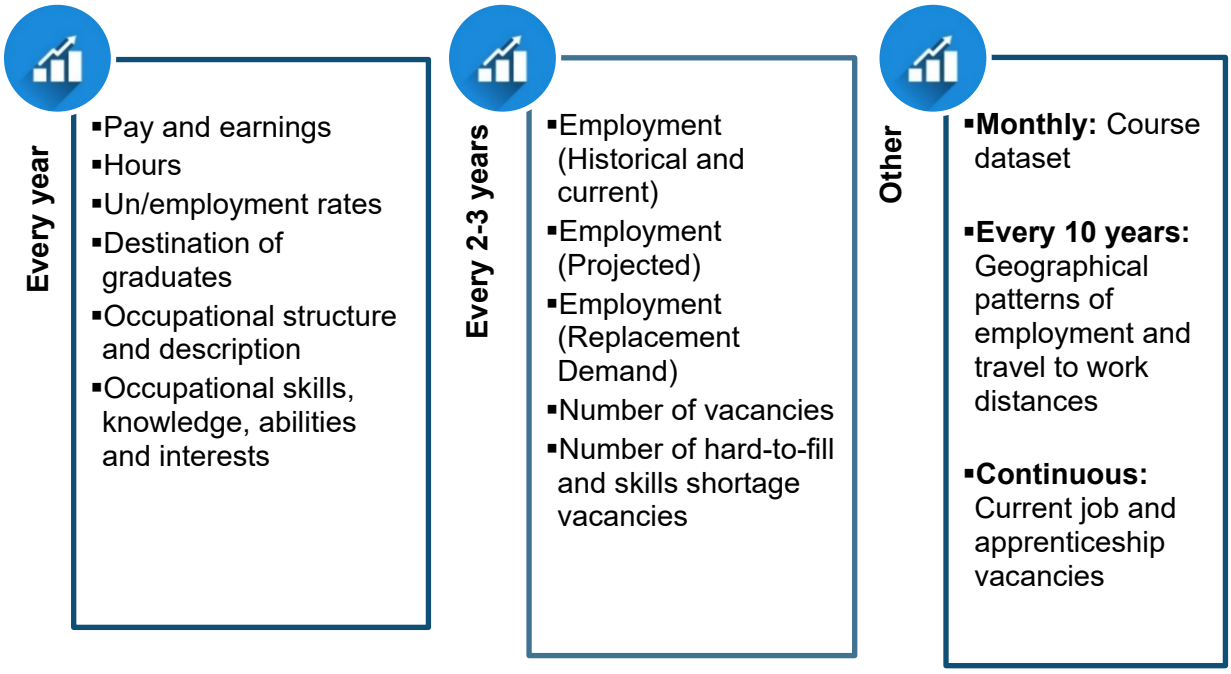
Due to changes in data delivery and the impact of COVID-19, the following datasets have not been refreshed during this reporting period:

- The pay and earnings (estimates based on ASHE and LFS) and hours (ASHE) datasets have not been refreshed. Access to data from the Secure Data Service is needed to undertake the econometric analysis to produce the pay and hours datasets. Data are accessed from a secure PC at the University of Warwick, but COVID-19 restrictions have meant that the data team have been unable to access the secure data. Revised secure data access procedures have recently been introduced by the UK Data Service enabling this work to be undertaken.

- In previous years, data from the Higher Education Statistics Agency (HESA) on higher education destinations have been obtained, processed and added to the database. However, changes to the HESA survey means that new data will be integrated into the service in 2021. An assessment of the new HESA data from the Graduate Outcomes survey was detailed in the 2018-2019 report.
- The number of vacancies, hard-to-fill and skills shortage vacancies provided by the Employer Skills Survey will be updated from the 2019 survey. It is expected that the 2019 data will become available shortly.

Figure 3.2, below, details the cycle for updating the data in LMI for All, including those datasets to be added in the next year. The next cycle will need to begin to deal with the switch from SOC2010 to SOC2020. This will have significant implications for the scale of work required both for data preparation as well as redevelopment of the main LMI for All API.

Figure 3.2 Overview of data update cycle and new data coming online



N.B. These timescales are based on previous update cycles.

3.3.1. New mappings

In the past, apprenticeship vacancy data were provided through the LMI for All service using the Apprenticeship service API, but this was withdrawn as Apprenticeship Frameworks have been incrementally replaced by Apprenticeship Standards since 2017. Instead, a mapping to UKSOC has been created in order for apprenticeship vacancies to

be made available through LMI for All. Each apprenticeship has been mapped to multiple SOC2010 codes where appropriate, as well as SOC2020 as this will be required in the future. This mapping has been reviewed and checked by experts in the field. A Beta version of Apprenticeship service vacancies API was released earlier this year, so the team are working to make these data available through the LMI for All service.

As part of the O*NET dataset, the service has included an O*NET to UK SOC2010 look-up table. This has, in the past, been made available on request to those researching and using O*NET data in the UK. During the last year, this has been made available through the LMI for All API and users can employ the look-up table for O*NET to UK SOC2010, and the reverse.

3.3.2. Course data

The Education and Skills Funding Agency (ESFA) have released, through the Open Government Licence, a course directory (comprising three datasets: courses with venues, providers, and regions). This directory has information on courses offered by over 4,000 learning providers who are contracted with the ESFA, including colleges, schools, private training providers and local education authorities. Courses include adult further education, adult community learning, apprenticeships, and courses aimed at people aged 16 to 18 years.

As this course directory is to be updated on a monthly basis, it will also be refreshed in LMI for All monthly. An ETL has been developed to ensure monthly updates are efficient and to check the structure and integrity of the dataset. The API will be updated to include this dataset, which has been fully indexed and a search interface is currently being developed. This will enable searches by course names and filtered by location of course (by town name, or county), and vice versa. This would return the following data:

- Course name;
- Course description;
- Start date;
- Cost;
- Location address;
- Provider;
- Course URL (where available);
- Attendance type (daytime, evening, weekend, day/block release);
- Delivery mode (classroom based, online, work-based); and

- Duration (number of hours, days, weeks, months, years).

Work in the coming months will focus on mapping location of course to county and/or Local Enterprise Partnership (LEP) area to complement other datasets in LMI for All.

3.3.3. Vacancy dataset

As part of developing and extending the data available through LMI for All, the team have piloted a web-scraping technique to collect UK vacancy data from online job portals.[6] Vacancy information is available from a range of different job portals. Some may display vacancies for a specific sector (e.g. NHS jobs for healthcare related occupations) or for different sectors (e.g. Guardian jobs). The collection period started in February 2020 to date for the Guardian and Reed, while for the NHS careers job portal the collection period began in May 2020. In order to collect as much data as possible on UK labour demand, it is necessary to collect, clean, standardise and compile vacancy data from different job portals.

The LMI for All service since it began has used vacancy data from a government vacancy stream. There have been issues with the structure, quality and consistency of these data, but significantly it only represented a small number of vacancies available in the UK. The aim of web-scraping job portals is to provide a constant stream of updated vacancy information in the LMI for All database to be made available through the API. By updating this dataset regularly (on a monthly basis during the pilot phase), it is hoped to create a close to real-time set of vacancy data.

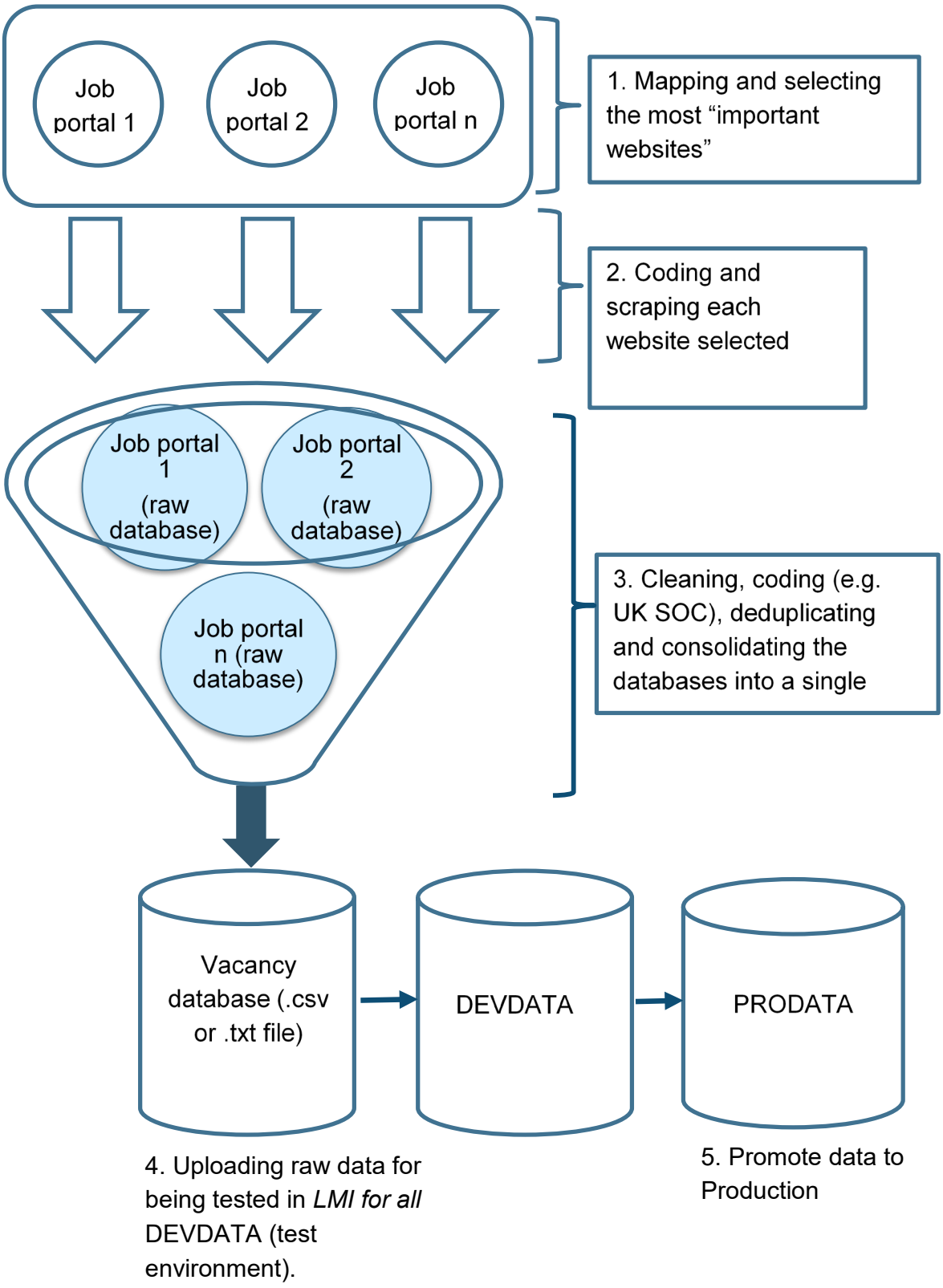
A pilot dataset was created scraping two job portals (The Guardian and Reed). This pilot dataset was coded to UK SOC2010 4-digit level, cleaned and quality checked by IER. It was then tested by the technical team to ensure it would work within the LMI for All database. Based on this testing, work is underway to create vacancy data from four major job portals in the UK (The Guardian, NHS careers, Reed and Jobs.ac.uk). These job portals are defined as significant in terms of the number of visits per day and volume of information available. The methodology for creating this vacancy dataset is detailed next.

Constructing a vacancy dataset: Methodology

Figure 3.3 summarises the general steps to build and integrate a vacancy dataset (from different UK job portals) compatible with the LMI for All infrastructure.

[6] 'Web scraping' consists of a computerised method to automatically collect information from across the internet; in this instance job portals are being scraped to collect information on advertised vacancies.

Figure 3.3 Steps for integrating the job portals information with the LMI for All infrastructure



First, job portals in the UK were identified and reviewed. Portals were selected if they provided a high volume of consistent information and represented a range of vacancies in the UK labour market. Four main elements were considered when selecting the job portals for the new dataset:

1. Volume (number of advertised job postings and number of job placements);
2. Website quality (detailed variables);
3. Traffic ranking (number of users as a measure of reliability);
4. Variety (cross-sector or cross-occupation information) (Cardenas, 2020; CEDEFOP, 2019).

To an extent, these four elements enable some representativeness and help ensure the quality of the vacancy dataset being created. After checking terms and conditions, the "best" sources of labour demand information were identified as The Guardian, NHS careers, Reed and Jobs.ac.uk.

Second, a web scraping code was developed for each job portal selected in order to automatically collect the relevant information (e.g. job titles, job descriptions, etc.) for each job posting. Each job portal needed a web-scraping code that recognised its unique (HTML - HyperText Markup Language) structure. As more job portals are selected more computational capability and manual coding is required.

The web scraping algorithms collect and store the vacancy information from each job portal. They were developed in a programming language, R. Once programmed, the job portal information can be regularly (e.g. daily) interrogated. A raw data file is generated for each job portal each time the code is run. A periodical manual check is undertaken to ensure that the algorithms are working correctly.[7]

Once the information is collected, the next step is to compile, clean, code (e.g. to UK SOC) and de-duplicate the job postings.

Next, each job portal dataset was cleaned, organised and categorised to consolidate into a single and cleaned vacancy database. These raw data are collected from different structures (e.g. number for variables, categories, etc.), so different programmes were developed to standardise and compile the different raw data into a vacancy dataset. These programmes use text-mining techniques (such as fuzzy merge, patterns identification, etc.) to standardise variables. For instance, to code the locality variable into

[7] Websites might change over time and thus, the web scraping codes need to be adjusted accordingly to those changes. The time to adapt a code depends on the number of changes in each website.

Local Enterprise Partnership (LEP) and county codes, a programme merges the locality names provided by employers in the job postings and the LEP and county classifications released by the ONS[8]. To some extent the programme deals with the fact that employers may use boroughs, wards, and cities among other names to indicate the location of the vacancy. Two steps were adopted to classify job titles for the LMI for All dataset. Initially, CASCOT was used to classify job postings to UK SOC2010 (see Jones and Elias, 2004). Then, it was necessary to complement this with machine learning techniques (see Cardenas, 2020) in order to code the maximum number of job postings to UK SOC2010 4-digit. An important feature of the vacancy dataset being created is that it is coded using UK SOC2010, which will enable it to be integrated into the LMI for All database. The software will have to be adapted to code vacancies to SOC2020 when LMI for All switches to the new SOC framework.

Once the information was standardised and coded, the raw datasets were merged into a single vacancy dataset. However, employers may publish the same vacancy on different job portals, so there was a process of deduplication. This involved scrutiny of each observation in terms of the values (such as experience, educational requirements, type of contract, localisation, year-month (when the vacancy was posted), wage, occupation and SIC) to identify duplicate job postings. It is worth noting that (so far) there is not a way to guarantee that all duplicate observations have been dropped through an automated or manual process. As shown by Cardenas (2020), when this process is implemented, the possible duplication issue does not considerably affect the validity of the vacancy database.

Finally, a visual inspection of the vacancy database was conducted to identify cases where the programming misclassifies a certain category (e.g. occupations, LEPs, etc.). Those misclassified cases were manually corrected and new programming produced for future data cleaning and coding. As a result, a single vacancy dataset (raw data) was generated. Table 3.2 lists the variables included in the dataset. The total number of observations (job postings) in the database are around 1.4 million for February to June 2020.[9]

[8] See: https://geoportal.statistics.gov.uk/

[9] The sample period for this pilot runs from 1 February 2019 to 11 June 2020, when the number of vacancies was probably depressed by the first COVID-19 ‘lockdown’.

Table 3.2 Structure of vacancy dataset under construction as of September 2020

Variable	Definition	Percentage of missing values
SOC Code	SOC Code (4 digits)	20.6% of job advertisements with missing values.
Job title	Short description of the job title offered	No missing values (mandatory field in the job advertisement).
Job link	Hyperlink to vacancy on job portal	No missing values (mandatory field in the job advertisement)[10].
Vacancy description	Detailed information about the profile required to fill the vacancy	No missing values (mandatory field in the job advertisement).
Location	Place where the vacancy is available, coded to county and LEP level	Around 18.8% of job advertisements with missing values for the county variable and 21.7% for the LEP variable.
Skills	Set of dummy variables that identify the skills required by employers according to ESCO	No missing values (this variable takes a value of 0 if a vacancy does not say anything related to skills).
Educational requirements	Set of dummy variables that identify the educational attainment required to fill the vacancy	Around 79.1% of job advertisements with missing values.
Labour experience	It takes values of 0 if the vacancy (explicitly) does not require any employment experience, it takes values of 1 if the vacancy (explicitly) requires any employment experience and it takes values of 2 if the vacancy does not mention anything regarding the employment experience required	No missing values (this variable takes a value of 2 if a vacancy does not say anything related to employment experience. Around 93.7% of job advertisements with missing values.
Wage	Continuous variable which indicates the amount of money that the hired person will receive	Around 25.2% of job advertisements with missing values.
Sector SIC	ISIC Code (at least 1-digit level)	Around 22.5% of job advertisements with missing values.
Job openings	The number of positions open for the vacancy	No missing values.

[10] The job links (URLs) collection started in late March 2020 onwards.

In the first instance, this vacancy dataset was uploaded to the LMI for All DEVDATA test environment in order to test new functionalities and the integrity of the new data. Once the tests were conducted, the vacancy dataset was transferred to the LMI for All production environment (LMI for All PRODATA server). Once fully integrated, the dataset will be refreshed on a monthly basis and checked on a weekly basis. These data will be provided alongside the Department for Work and Pensions' Findajob API data that is currently provided through LMI for All and mapped to SOC using a fuzzy matching process.

In the next stage of development, the following will be undertaken to improve the overall quality of data provided: further cleaning of the data, reviewing missing observations (see Table 3.2) and adjusting algorithms and coding to significantly reduce missing observations.

3.4. Implications of the change from SOC2010 to SOC2020

The LMI for All team have been preparing for the implementation of UK SOC2020[11] and how the service will need to move to this new organising framework, which is currently based on UK SOC2010. ONS have now introduced SOC2020 which will take over as the main way of classifying occupations over the coming years. In this classification there have been significant changes at the SOC minor and unit group levels, which reflect the changing structure of occupations in the UK. Some unit groups have been disaggregated into less heterogenous groups and, where occupations have been declining in size, some unit groups have been aggregated.

Updating to SOC2020 will entail the following tasks for the LMI for All service:

1. Reprogramming to base the service on the new categories;
2. Collection and uploading of new data using the new classification; and
3. Development of historical information based on the new classification.

This work will require more time and resource in comparison to the recent cycle of annual updates, particularly around the development of historical information based on the new classification (task 3). The *Working Futures* database provides a foundation for much of this work and the historical data included will need to be updated to the new classification before task 3 can be undertaken. It is likely that the *Working Futures* database will

[11] From more information see: https://www.ons.gov.uk/methodology/classificationsandstandards/standardoccupationalclassificationsoc/soc2020

continue to rely on the old classification, so there will be a need to translate historical data and the projections on to the new SOC2020.

Furthermore, in order to do the updates, a number of datasets (such as the LFS and ASHE) would need to be using the new classification in order for the LMI for All service to make the change. Once LFS data are available with SOC2020, a method of incorporating this new information into the database will be required. This could include either: translating the information back into the old SOC2010 categories; or reprogramming the service to use the new categories based on SOC2020. These options are being reviewed and considered by the team. The shift to UK SOC2020 for the LMI for All service will be required to continue meeting users' needs. To achieve this shift, work needs to be undertaken to update datasets, restructure the LMI for all database, and reprogramme the API. The technical infrastructure of LMI for All will, however, enable work to be undertaken in the development data server before it is moved to the production environment. This will be necessary to ensure continuity of service and to minimise the impact on user applications and websites.

3.5. Summary and next steps

Over the reporting year, work has been undertaken to refresh existing datasets and create new mappings to support the use of the data. Whilst COVID-19 restrictions have delayed the release of the updated pay, earnings and hours datasets, the econometric model has been prepared and all programming updated using R (software environment for statistical computing). The development of the job vacancy dataset created by web scraping has taken a significant amount of time, but this will be an enhancement to the service and provide some local level data currently missing from LMI for All.

A major piece of work on the LMI for All service in the coming year will be managing the change to UK SOC2020. The data and technical teams will be working closely on implementing this change, which will require: restructuring the database; major reprogramming of the API; the collection and uploading of new data using the new classification; and the development of historical information in the database based on the new classification. It will be a priority to maintain continuity of service and ensure the changes are communicated.

4. Stakeholder engagement

The stakeholder engagement strand of the project continues to focus on awareness raising of the potential of the LMI for All service amongst a range of relevant stakeholder groups and audiences. A key message has focused on how the service can provide information to support individuals to make informed decisions about learning and work. Key stakeholders have remained careers service providers, education institutions, training providers, local governments and agencies, developers, and parents and carers. The objectives for this phase are similar to previous phases of the project, informing and shaping activities. However, it should be noted that these objectives were agreed before the onset of the COVID-19 pandemic, so outcomes necessarily reflect the new circumstances for delivery. Delivery of stakeholder engagement activities has mostly been online over the last year.

4.1. Objectives and outcomes

As intimated above, varied stakeholder activities were agreed for this phase of work in October 2019, framed by a stakeholder plan that was developed in consultation with the Department for Education at the start of the project. Many of the stakeholder activities addressed several objectives. These objectives and how they have been addressed is set out below.

- To facilitate and organise events and workshops to promote LMI for All to key stakeholders. The team had planned to organise a number of targeted workshops, webinars, training events and presentations to inform and update stakeholders about LMI for All. However, given the new circumstances imposed by the pandemic, which prevented face-to-face contact, there has been a heavy reliance on activities undertaken remotely and online.
- To participate in third party events, which target relevant audiences. The team has continued to contribute to various events to reach a range of stakeholders, such as open data specialists, developers, policy makers, career practitioners and researchers. The aim has been to promote the LMI for All service, its current third party users, Careerometer and Skillsometer. Again, because of the pandemic, the methods of participation have been limited to online, distanced events for different audiences.
- To continue negotiations with priority stakeholders to support them in the development of applications that use LMI for All. Designers and developers of websites, applications and widgets using LMI for All have been engaged in informal conversations and provided support where requested, online.

- To facilitate user fora and/regular meetings involving organisations that draw on LMI for All, to gather feedback on support needs and potential improvements to the service and gain insights into who end-users are, and how they use the data and the benefits in terms of careers decision-making. This has been achieved through social media activities, blogs, webinars and meetings with developers.
- To support the Department for Education in understanding what more government can do to support the development and use of high-quality careers apps and websites, including increasing awareness amongst schools and colleges of tools that exist. The team have delivered three presentations to different groups within the Department for Education to increase understanding and awareness of the functionality and potential of LMI for All, as well as information on the work to expand the database and service.
- To undertake routine evaluation of activities in a manner appropriate to each particular activity and target group. Wherever possible, evaluations have been undertaken and positive feedback received.

These objectives were achieved through integrated activities of work, which span this reporting year until the end of the project. Activities have included the enhancement of the LMI for All website (see section 2.5), the publication of blogs on the website and social media activities.

Like the previous phase, a consistent end-user focus has been retained in these interrelated work clusters. The end-users comprise the third parties that have continued to be the data and service users, together with potential clients and customers of those end-users. Whilst the main geographical target for these stakeholder activities continued to be the UK, the methodology underpinning the initiative has attracted international interest. Because of restricted travel during the pandemic, webinars have been the main method of dissemination during this reporting year. Invitations were received from the professional associations for both career practitioners and career teachers in New Zealand, Canada and Australia where there is a great deal of interest in the development process, the technological challenges and the need to integrate LMI into all aspects of career practice.

4.2. Stakeholder plan

As in previous years, the development of the stakeholder plan was the first activity for this strand, conducted in close consultation with the Department for Education. Objectives for the stakeholder plan reinforced and reflected those of previous years as listed below:

- To raise awareness and understanding of the LMI for All service;

- To support third party development of applications and/or enhance existing applications so that the reach of LMI for All among end-users was increased;
- To ensure the implementation and use of Careerometer;
- To ensure the development and implementation of Skillsometer, to complement the use of Careerometer, amongst end-users;
- To drive traffic to existing applications of LMI for All; and
- To foster positive and supportive attitudes to the LMI for All service among target stakeholder groups.

This plan included a broad forward schedule for activities and key stakeholder audiences. Target stakeholders were agreed as: schools; career hubs; further and higher education providers; careers professionals and careers organisations; web and app developers; information and data service providers; third party users (defined as those using LMI for All in their online service); the Department for Education, local authorities and other government departments; and parents and carers.

The protocol (agreed with the Department for Education and implemented last year) for vetting and approving specific activities identified by the project team as suitable was continued. So, for example, whilst presentations at national and international conferences was agreed broadly as a suitable stakeholder activity, whenever a specific conference was identified, permission was sought from the Department for Education, providing details of the conference, audience, costs, location, etc. Added value to the project was demonstrated by some events where invitations to present had meant costs were covered elsewhere, such as by hosts.

For each event, audience/participant numbers were systematically recorded, subsequent to the activity, together with the exact nature of the activity carried out and any follow-up that had arisen. For online events (like webinars) feedback was noted, particularly where there was an indication of how the presentation could be improved (e.g. like providing links to publications), though this feedback was difficult to collect verbatim. A short report was also presented by the relevant project team member at the monthly management team meeting with the Department for Education.

4.3. Stakeholder activities and dissemination of the LMI for All service

Reflective questions related to stakeholder activities and posed by the Department for Education were at the forefront of planning for this reporting year and included the following:

- Are the activities spread evenly throughout our target groups?
- Are there a range of activities?
- Are there any additional stakeholders we want to engage?
- Are the proposed activities aligned to the stakeholder engagement objectives?

The design of the stakeholder strategy sought to address these questions. The COVID-19 pandemic, however, resulted in the postponement of seven planned stakeholder events that would have helped address any imbalance of target groups, for example, those working in Further Education. Despite these postponements, LMI for All was promoted at 13 dissemination events, ranging from national UK events and international conferences, workshops and presentations to local and regional events. In addition, ten targeted events and meetings were undertaken with over 30 individuals, developers and organisational representatives. These were follow-ups from activities as well as engagement with specific queries about the service. The number of stakeholders reached directly through these events is an estimated 712. In addition, the Canadian professional association invited the submission of a blog, which was circulated to an audience of 11,500. The six international events in which LMI for All was presented, illustrated and confirmed yet again the innovative nature of the service for careers provision and how this continues to stimulate interest in the broad community of career practice.

As in the previous year, three workshops on LMI and LMI for All were delivered for the Career Development Institute reaching around 40 participants. The purpose of the events was to support careers practitioners with their professional development, namely the identification, application and assessment of labour market information and intelligence. These workshops enable LMI for All as a source of high quality LMI to be presented alongside LMI sources, many of whom are third party users of LMI for All. These events have consistently received positive feedback and generated interest in Careerometer and those websites and applications that use the LMI for All service.

Building on the positive feedback received in the previous year, the approach adopted to disseminate events went beyond presenting a factual account of LMI for All's development, by examining the impact of this type of ICT development on the professional identity and practice of career and employment practitioners. This approach continued to be very well received by mixed audiences of practitioners, managers, trainers, policy makers and researchers. Findings from other recent, relevant research were built into sessions as this provided context and evidence for the value of LMI. For example, findings from research into the features of lifelong guidance policy and practice across Europe, included LMI and the use of technology in careers as key features (Barnes, Bimrose, Brown, Kettunen, and Vuorinen, 2020). Findings for this research, which was managed and conducted by the IER team working on LMI for All, highlighted

the importance of LMI to efficacious lifelong guidance systems. Related to this finding, the LMI for All initiative was highlighted by participants in the European research as an innovative practice. It was therefore detailed in the report for the European Commission and subsequently highlighted at the European vocational education and training conference in Finland. Overall, four webinar presentations (approximately 140 participants) were delivered that highlighted and discussed the finding related to LMI and LMI for All, including a three day international summer school (online), hosted by Jönköping University, Sweden, which 29 doctoral students, research academics and tutors attended.

One other research study also carried out by IER, into supporting parents and carers in providing career support to young people in their care found how LMI was valued by parents and carers, but it was difficult for them to access (Barnes, Bimrose, Brown, Gough, and Wright, 2020). It also highlighted how practitioners engaging parents and carers in careers activities needed access to high quality, reliable labour market information, so that they had greater awareness of labour market changes and opportunities. Again, this finding was disseminated and highlighted at six webinars (approximately 300 participants). Generally, the more that evidence regarding the use of LMI in broader aspects of career practice was integrated into dissemination events, the more positive the feedback on the session delivered by the team, though this feedback remained at a general level, rather than focusing on particular aspects of the session.

It has been found in previous years that detailed evaluation of stakeholder activities is not always possible at all events, especially as some events are organised by a host who undertakes their own evaluation. Webinars are even more difficult to formally evaluate, since they have all been organised by hosts. However, the project team were able to assess the value of stakeholder activities based on interest expressed in terms of the number of participants signing up to sessions, and then attending, as well as questions and requests for follow-up webinars, and/or information during the activity.

In terms of what worked well for stakeholder engagement during the current project phase, ongoing dialogue continues to emerge as critically important. The pandemic, however, compromised the ability to address all the reflective questions posed by the Department for Education at the beginning of this section, particularly the need to ensure activities were spread evenly across target groups.

4.4. Engagement through the LMI for All website and social media

The LMI for All twitter account is managed by the LMI for All team. Over the reporting year, there has been between 4-6 tweets per month that focus on LMI for All news,

stakeholder activities and promotion of the service and widgets. The account follows those that use LMI for All and others engaged in LMI work and research. Examining the Twitter account analytics, the following summarises activity of the last year:

- 34,915 Tweet impressions (this is the number of times users on Twitter have viewed an LMI for All tweet in their timeline);
- 169 New followers (currently, LMI for All has 1,355 followers);
- 64 mentions (this is the number of times others have mentioned LMI for All using the Twitter handle);
- 1,005 Profile visits (this is the number of times the LMI for All Twitter profile page has been visited).

On the LMI for All website, eight blogs[12] have been published covering a range of topics including the following:

- Piloting a new online job-scraping tool;
- Can Big Data fill your data vacuum?;
- Labour market and skills projections: 2017 to 2027;
- Professional identity transformation: supporting career and employment practitioners at a distance;
- Understanding the gender pay gap;
- LMI for All team meets Studiekeuze123;
- Working Better: Using data for inclusive system for jobs; and
- Jordan British Council and AQACHEI visit LMI for All team.

4.5. Supporting careers practitioners

Career practitioners and their managers remain a key target group for stakeholder activity. This has been achieved through the delivery of professional development events and the development of online LMI learning units on the LMI for All website.

An effective way to reach careers practitioners across a range of sectors has continued to be offering collaborative events with the Career Development Institute (CDI). Up to this year, the CDI have taken the lead in organising events and the team have developed materials and delivered the events in person. However, this traditional method of

[12] For all the blogs on the LMI for All website see https://www.lmiforall.org.uk/news-and-blog/

connecting with the broad community of career guidance practice was compromised by the pandemic. The CDI went online and the virtual offers of training support to membership (c 5,000 members) were broadened. The LMI for All team offered webinars on LMI for All, which were taken up with enthusiasm. In total, four webinars were delivered to CDI membership. For each event, the maximum number of participants (100) was registered, with a total of 283 attending and participating in webinars. A repeat of the webinar that focused on LMI for All as an integral part of practice, was requested. This included promotion of the seven LMI learning units (see below). Positive comments about the webinars overall related to relevance and practicality. This suggests that online events remotely could be as effective as face-to-face dissemination events in supporting careers practitioners with their understanding and effective use of LMI in practice.

Seven short learning units have also been developed by the LMI for All project team to support career practitioners and teachers using or interested in accessing LMI as part of the careers education, advice, guidance and counselling process. These materials are based on the training sessions run by the team in collaboration with the Career Development Institute (CDI).

The learning units have been designed to help users understand the challenges and issues with accessing and using LMI, plus other complex aspects of LMI, in more depth and breadth. The aim is to support users in their understanding and awareness of the range of different datasets and variables available in the LMI for All database. Each unit provides learning objectives, some basic material on the unit topic, some relevant resources and questions on which to reflect. The seven learning units cover the following topics:

1. What is LMI and why is it important?
2. Who uses LMI and what for?
3. Sources of LMI
4. Limitations of LMI
5. Features of LMI
6. Choosing amongst sources of LMI
7. Understanding replacement demand and skills shortages

Currently, these units have been added to the ‘Explore LMI’ section of the LMI for All website, but the intention is to extend these units and create a new section on the website for online learning materials to be located.

4.6. Take-up of LMI for All and Careerometer

The LMI for All service continues to be used by a range of users within their websites and learning platforms (see also section 2.3). Many of the users continue to develop their offering by drawing upon more data from LMI for All, including Eluceo, U-Explore's Start, BBC Bitesize Careers and the National Careers Service. The LMI for All team are currently working to update the case studies on the LMI for All website and develop new case studies with third party users. These case studies provide useful insights into how the LMI for All service can be used.

Careerometer continues to be embedded in a number of careers and learning providers' websites and learning platforms. Whilst it is becoming more difficult to calculate the number of sites with Careerometer, there are over 160 examples accessible on the internet. Skillsometer was only released in August 2020, but it is already being used. The take-up of Skillsometer will be monitored and reported on in the next report.

4.7. Summary and next steps

As in previous years, there continues to be a substantial and sustained uptake of the LMI for All resource. The ambition is to expand and secure this sustained uptake. Even during this wholly exceptional year of the COVID-19 pandemic, the strategy relating to integrated stakeholder activities has provided a clear demonstration of a consistent focus on end-user need. It also continues to represent a transparent engagement strategy with those third parties that are the primary data/service users. The overall aims of these activities are to stimulate demand and promote the data tool, with the evaluation evidence that is available indicating that these aims are being met.

The stakeholder activities were probably less diverse than in previous phases of the project because the restrictions imposed by the pandemic brought a heavy reliance on remote working. They have all, however, without exception, been aligned with the stakeholder engagement objectives and plan. There are some stakeholders that have not yet been effectively engaged, which could be the focus of further effort, specifically parents/carers (some successful work has been completed), further education colleges; and career teachers/leaders. Activities that targeted these audiences were planned, but had to be postponed. Over the next year, stakeholder activity will focus on:

- Increasing activities relating to targeting and awareness raising in schools and further education colleges, targeting careers leaders as well as parents and carers;
- Continuing work with those stakeholders who have started designing and developing LMI for All interfaces and applications.

Reflecting on developments for the next stage of the project, effectively supporting practitioners in the use of LMI to stimulate greater engagement with LMI for All continues to be a focus going forwards. Delivering webinars is one way that has been very successfully implemented. Attention needs to be given on how best to enhance and support the confidence and competence of career and employment practitioners operating in compulsory and tertiary education.

Over the next year, a range of stakeholder activities will continue to be organised as well as further development and refinement of case studies that demonstrate how third party developers have used LMI for All and the benefits accrued. The team will also explore how to promote the service through various publications, including its own newsletter. Currently, approximately 350 people have registered an interest in receiving further information on the service.

5. Future developments and next steps

Stakeholder engagement, technical and data developments outlined in this section will be undertaken over 2020-2021 and progress reported in the next annual LMI for All report. The following details the future objectives for each strand of the LMI for All project ensuring that the service is well maintained, continues to be useful and relevant, addresses users' (both career service providers and their clients/customers) needs, and importantly takes advantage of IT and data science advancements.

Technical developments

1. To continue development of the ETL system

   Since this provides another layer of data checks, it will safeguard data quality as well as ensuring that data transfers from the data team to the technical team is more efficient. The ultimate beneficiary will be the end user, in receipt of a higher quality product in the form of LMI.

2. To continue to maintain and monitor the LMI for All infrastructure, including test and production servers and provision for backup services ensuring that the data made available to developers and end-users meets required quality standards.

   Crucially, this will ensure continuity of service, together with an efficient service. It also contributes towards helping to maintain data quality. Continuity of service is essential to maintain end-user confidence in the LMI for All service, providing reliable and current LMI for end-users.

3. To update the infrastructure to enable access to new datasets and monitor it in terms of operation.

   This will provide considerable added value to the service, by addressing the data needs of a range of users. Without an effective and efficient infrastructure, the service would cease to be an attractive source of LMI to end users.

4. To prepare the LMI for All infrastructure for UK SOC2020.

   This future-proofing of the service, which improves and develops the infrastructure on which the service is built, will ultimately enable users to access data that are relevant to them and reflective of the current labour market structure. Thus, this will enhance the quality of data service provided to end-users.

5. To continue enhancements to the LMI for All website and service.

   Since the website and service represent the primary source of communication with a range of stakeholders, continued enhancement comprises added value by increasing the quality of this mode of communication. As the range of stakeholders is gradually being increased, as well as the overall service offered, clear and

effective communication becomes more critical, since it is the channel through which end-users are able to communicate with the LMI for All team.

Data developments

1. To continue refreshing existing datasets held in LMI for All and management of documentation for each dataset.

   Pivotal to the currency and legitimacy of the service, this is key to the service being of value to users because the process ensures that data are up-to-date. A key objective of the service has always been about the quality of data, which implicates the highest levels of currency achievable.

2. To review quality assurance mechanisms and check data rendered by the API are consistent with values contained in the dataset generated.

   Providing data that are of high quality and safeguarding disclosive or unreliable data are key characteristics of the service. Without the ability to claim high quality, reliable and current data, the service would not be credible.

3. To review and implement new datasets to expand the LMI for All service.

   Horizon scanning for potential additional datasets adds value to the service because the broader the scope of high quality current data available through the service, the more likely it is to have wider appeal to help third party users in the support they provide to their clients and customers.

4. To prepare for the move from UK SOC2010 organising framework to UK SOC2020.

   This will be a significant part of developing the LMI for All service to ensure it reflects the updated occupational structure of the current labour market. SOC2020 will be a critical framework upon which the database is organised and delivered to end-users of the service.

Stakeholder engagement activities

1. To continue to stimulate demand and promote the LMI for All service to a range of stakeholders, with a focus on careers hubs, further education institutions, local governmental organisations and open data specialists.

   The gradual expansion of the appeal of the service to a wider constituency of end user groups continues to be a key objective of the LMI for All project, because this delivers maximum impact on an increasingly broad range of stakeholders. High quality career practice requires recourse to high quality and reliable LMI data. Promoting LMI for All across as many members of the broad career practice community as feasible is essential to emphasise the importance of integrating LMI into career decision making. It remains the case that the continuing professional

development available for career practitioners and their managers would benefit from external support in using LMI effectively in practice.

2. To maintain support and dialogue with existing third-party users of LMI for All and continue working with those stakeholders who have started designing and developing LMI for All interfaces and applications.

   Maintaining a consistently rapid and high quality response to third party users who are on the brink of engaging with the LMI for All service is essential to increase and support accessibility to the service. It also plays an important role in contributing to meeting both the technical and data needs of these users.

3. To increase activities relating to the targeting and awareness raising in schools, targeting careers leaders as well as parents and carers, making full use of ICT.

   Available evidence confirms the influence exerted by parents and carers on the career development of young people in their care. This group, together with career leaders, need and want to have access to high quality LMI so that the support they provide young people is more likely to be efficacious. Increasing activities in this area will assist with maintaining existing communication channels with key stakeholder groups, whilst the LMI for All team learn and adapt to new ways of working due to COVID-19.

4. To continue collaboration with organisations that support the professional development of careers practitioners to ensure practitioners are confident in their use of LMI and in choosing between sources of LMI.

   Professional, and similar, organisations are crucial intermediaries in the process of supporting those that use LMI in the career guidance services and provision to end-users, for example, by providing access to continuing professional development (CDP). They can also provide an invaluable role in promoting the service to members. Consequently, continued collaboration with such organisations continues to be a key part of the LMI for All service and part of the add-value the service provides.

References

Alexander, R., McCabe, G. and De Backer, M. (2019). *Careers and labour market information: An international review of the evidence*. Reading: Education Development Trust. Available from: https://www.educationdevelopmenttrust.com/EducationDevelopmentTrust/files/3a/3a1edb39-3ce4-46fb-958d-6e4f2fe3bc78.pdf

Barnes, A. (2015). *Ten Best Resources for LMI*. Cegnet, Available from: http://www.cegnet.co.uk/uploads/resources/The_ten_best_LMI_resources_final.pdf

Barnes, S-A., Bimrose, J., Brown, A., Gough, J. and Wright, S. (2020). *The role of parents and carers in providing careers guidance and how they can be better supported: International evidence report*. Coventry: University of Warwick. Available from: https://warwick.ac.uk/fac/soc/ier/research/careerguidanceparents/ier_gatsby_jpm_parents-carers_final_report_v8_final.pdf

Barnes, S-A., Bimrose, J., Brown, A., Kettunen, J., and Vuorinen, R. (2020). *Lifelong guidance policy and practice in the EU: trends, challenges and opportunities*. Luxembourg: Publications Office of the European Union. Available from: https://ec.europa.eu/social/main.jsp?catId=738&langId=en&pubId=8284&furtherPubs=yes

Cardenas, J. (2020). *A web-based approach to measure skill mismatches and skills profiles for a developing country: the case of Colombia* (Unpublished doctoral dissertation). University of Warwick, Coventry, United Kingdom.

CEDEFOP (2016). *Labour market information in lifelong guidance*. Thessaloniki, Greece: European Centre for the Development of Vocational Training. Available from: http://www.cedefop.europa.eu/files/5555_en_case_study_england.pdf

CEDEFOP (2019). *Online job vacancies and skills analysis: a CEDEFOP pan-European approach*. Thessaloniki, Greece: European Centre for the Development of Vocational Training.

Department for Education (2017). *Careers Strategy: Making the most of everyone's skills and talents* (Policy paper). London. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/664319/Careers_strategy.pdf

Dickerson, A., Wilson, R., Kik, G., and Dhillon, D. (2012). *Developing Occupational Skills Profiles for the UK: A Feasibility Study.* UKCES Evidence report, 44. Wath-upon-Dearne: UK Commission for Employment and Skills.

Gatsby (2014). *Good Career Guidance*. London: The Gatsby Foundation. Available online: https://www.gatsby.org.uk/uploads/education/reports/pdf/gatsby-sir-john-holman-good-career-guidance-2014.pdf

Holland, J. L. (1997). *Making vocational choices: A theory of vocational personalities and work environments* (3rd ed.). Odessa: Psychological Assessment Resources.

Jones, R., and Elias, P. (2004). *CASCOT: Computer-assisted Structured Coding Tool*. Coventry: Institute for Employment Research, University of Warwick.

Mezzanzanica, M. and Mercorio, F. (2019) *Big Data for labour market intelligence: An introductory guide*. Torino, Italy: European Training Foundation. Available from: https://www.etf.europa.eu/sites/default/files/2019-06/Big%20data%20for%20LMI.pdf

OECD (2015). *OECD Reviews on Local Job Creation Employment and Skills Strategies in England, United Kingdom*. Paris: OECD Publishing. Available from: http://www.oecd.org/cfe/employment-and-skills-strategies-in-england-united-kingdom-9789264228078-en.htm

Sunak, R. (2020). *A Plan for Jobs 2020 speech as delivered by Chancellor Rishi Sunak*. HM Treasury, UK. Available from https://www.gov.uk/government/speeches/a-plan-for-jobs-speech

The Careers & Enterprise Company and WorldSkills UK (2016). *How to use labour market information*. London. Available from: https://www.worldskillsuk.org/media/4461/how-to-use-lmi-booklet-a5.pdf

Annex A: LMI for All reports

The following is a list of LMI for All reports which document the development of the service from piloting, to full project status, to its continued development and enhancement.

Table A.1 Overview of LMI for All development

Phase	Date	Activity
1	2012	Tested the feasibility and viability of the foundation concepts and ideas. Bimrose, J., Wilson, R., Elias, P., Barnes, S-A., Millar, P., Attwell, G., Elferink, R., Rustemeier, P., Beaven, R., Hay, G. and Dickerson, A. (2012). *LMI for All Career Database Project - Processes Adapted and Lesson Learned*. London: UKCES. Summary available to download at: http://www2.warwick.ac.uk/fac/soc/ier/research/prototypecareerslmi/bimrose_et_al_2012_lmi4all_summary.pdf
2A	2012-2013	Expanded and updated the database, tested and evaluated the API, improved the technical infrastructure, explored data sets that could add value to services, raised awareness of the service through a number of stakeholder events. Bimrose, J., Wilson, R., Barnes, S-A., Owen, D., Green, A., Li, Y., Millar, P., Bosworth, L., Holden, A., Attwell, G., Rustemeier, P., Elferink, R., McGreggor, A., Wicks, D., Long, A. and Dickerson, A. (2013). *LMI for All Developing a Careers LMI Database Phase 2A report.* London: UKCES. Available from: http://www2.warwick.ac.uk/fac/soc/ier/publications/2012/bimrose_et_al_2012_lmi4all_summary.pdf
2B	2013-2015	Continual development of the LMI for All service, implemented recommendations from Phase 2A. Bimrose, J., Wilson, R., Barnes, S-A., Owen, D., Li, Y., Green, A., Bosworth, L., Millar, P., Holden, A., Attwell, G., Rustemeier, P., Elferink, R. and Higginbottom, J. (2015). *LMI for All: Developing a Careers Database (Final report)*. London: UKCES. Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/446332/LMI_for_All_Final_report_v020715_FW.pdf

Phase	Date	Activity
3	2015-2016	Continual development of the LMI for All service, updated database and technical infrastructure, continued to raise awareness of the service through stakeholder events, worked with stakeholders to develop LMI for All applications. Bimrose, J., Sofroniou, N., Barnes, S-A., and Attwell, G. with Rustemeier, P., Stieglitz, D., Elferink, R., Hughes, D., Li, Y., Owen, D., Wilson, R., Bosworth, L. and Millar, P. (2016). *LMI for All: Technical, Stakeholder engagement and data development services*. Wath-upon-Dearne: UKCES. Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/545324/LMI_For_All_for_web.pdf
4	2016-2017	Continued development of the LMI for All service through a range of stakeholder engagement services and data refreshing, as well as the implementation of a new IT infrastructure to manage increasing queries to the database. Bimrose, J., Barnes, S-A., Owen, D., Hughes, D., Wilson, R., Attwell, G. and Rustemeier, P. (2018). *Labour market information (LMI) for all: Stakeholder Engagement and Usage, Data and Technical Developments*. London: Department for Education. Available from: https://www.gov.uk/government/publications/labour-market-information-lmi-2016-to-2017-data
5	2017-2018	Continued maintenance and development of the LMI for All service. Barnes, S-A., Bimrose, J., Owen, D., Wilson, R., Attwell, G. and Rustemeier, P. (2019). *Labour market information (LMI) for All 2017 to 2018*. London: Department for Education. Available from: https://www.gov.uk/government/publications/labour-market-information-lmi-for-all-2017-to-2018
6	2018-2019	Development and enhancement of the LMI for All service documenting the work undertaken in the first of a three year programme of work. Barnes, S-A., Bimrose, J., Owen, D., Wilson, R., Luchinskaya, D., Attwell, G. and Rustemeier, P. (2020). *Developing and enhancing a labour market information database: LMI for All Stakeholder Engagement and Usage, Data and Technical Developments (2018-2019)* (Research report). London: Department for Education. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/863703/LMI_for_All_2018_2019.pdf
7	2019-2020	This report describes the activities undertaken during this phase of the project.

Department for Education

© Department for Education 2021

Reference: RR1096

ISBN: 978-1-83870-233-5

The views expressed in this report are the authors' and do not necessarily reflect those of the Department for Education.

Any enquiries regarding this publication should be sent to us at:
LMI.DFE@education.gov.uk or www.education.gov.uk/contactus

This document is available for download at www.gov.uk/government/publications