LMI for All Data descriptions and indicators Release 19 June 2014

Table of Contents

Institute for Employment Research, University of Warwick, *Version 19 June 2014*

Summary of main datasets and indicators included in the LMI for All Database

The current version of the LMI for All database contains the following key data sets:

1. Employment, projected employment and replacement demands from *Working Futures*
2. Pay and earnings based on the Annual Survey of Hours and Earnings and the Labour Force Survey
3. Hours based on the Annual Survey of Hours and Earnings
4. Unemployment rates based on based on the *Labour Force Survey*
5. Skills shortage vacancies based on the *Employer Skills Survey*
6. Skills, Abilities and Interests based on the US *O*NET* system
7. Occupational descriptions from the ONS Standard Occupational Classifications
8. Current vacancies available from *Universal JobMatch*
9. Higher education destinations data from *HESA*

Dataset 1. Employment

Provides very detailed information on job opportunities which covers the following dimensions:

- detailed SOC2010 4 digit occupational categories;
- highest qualification held;
- industry;
- countries and English regions within the UK;
- gender; and
- employment status (full-time and part-time employees or self-employment).

As well as historical employment levels this also includes projected **Employment** levels to 2022 and estimates of **Replacement Demands** (which is a measure of job openings likely to arise over a selected period because of people currently employed retiring or leaving for other reasons). Please note data from 2000-2011 are only available at SOC 2010 2 digit level.

Dataset 2. Pay and earnings

Provides the same level of detail on **Pay** but does not include any projections. The current data are available for 2012-2013. It is planned to extend this as new data become available on the same basis. *Pay* data are available for the following dimensions:

- detailed SOC2010 4 digit occupational categories;
- highest qualification held;
- industry;

- countries and English regions within the UK;
- gender;
- age; and
- employment status (full-time and part-time employees or self-employment).

Average, median and decile pay data are available.

Dataset 3. Hours

Provides a similar level of detail for *Hours* but does not include any projections. Nor does it cover qualifications. The current data are restricted to 2012-2013. It is planned to extend this as new data become available on the same basis. *Hours* data are available for the following dimensions:

- detailed SOC2010 4 digit occupational categories;
- industry;
- countries and English regions within the UK;
- gender; and
- employment status (full-time and part-time employees or self-employment).

Dataset 4. Unemployment rates

Provides information on *Unemployment rates*. In principle a similar level of detail is provided as for Pay (again excluding any projections). *Unemployment rates* are available for the following dimensions:

- detailed SOC2010 4 digit occupational categories;
- industry;
- countries and English regions within the UK;
- gender; and
- employment status (full-time and part-time employees or self-employment).

However, small sample sizes mean that many of the more detailed results cannot be generated or published.

In practice some of the details here and elsewhere may need to be suppressed because of concerns about privacy, disclosure of commercially sensitive information and lack of statistical robustness. Details of this are given in the description of the individual data sets.

Dataset 5. Number of vacancies

Provides information on *Vacancies*. At present this is based on the Employer Skills Survey and focuses on skills shortage vacancies for 2011 and 2013. In principle, incorporation of a much broader range of vacancy information based on Job Centre data is planned, but this is not currently available in a suitable format.

Institute for Employment Research, University of Warwick, *Version 19 June 2014*

Dataset 6. Skills, abilities and interests

Provides detailed data on the *skills*, *abilities* and *interests* associated with particular occupations (SOC2010 4 digit categories). The data are taken from the well-respected US O*NET system. They are tied to the UK occupational categories by a mapping that links each SOC2010 4 digit occupation with one or more US occupations.

Dataset 7. Occupational descriptions

Provides a detailed structure of SOC 2010 occupations together with descriptions for each occupation. These data are from the ONS Standard Occupational Classification database.

Dataset 8. Current vacancies

Provides, through a fuzzy search, vacancies from Universal JobMatch. Matching these vacancies to SOC 2010 occupations in planned for 2014.

Data overview – LMI for All API release 19 June 2014

Dataset	Employment (Historical)	Employment (Projected)	Employment (Replacement Demand)	Pay and earnings	Hours	Unemployment Rates	Number of vacancies	Skills, Abilities, Interests	Occupational descriptions	Current vacancies
Source	Working Futures/ Business Register and Employment Survey (BRES)/ Labour Force Survey (LFS)	Working Futures/ Business Register and Employment Survey (BRES)/ Labour Force Survey (LFS)	Working Futures/ Business Register and Employment Survey (BRES)/ Labour Force Survey (LFS)	Annual Survey of Hours and Earnings (ASHE)/ Labour Force Survey (LFS)	Annual Survey of Hours and Earnings (ASHE)	Labour Force Survey (LFS)	Employer Skills Survey (ESS)	O*NET database	ONS Standard Occupational Classifications	Universal Jobmatch
Indicator	Number of jobs (employee, self-employed)	Number of jobs (employee, self-employed)	Number of jobs openings between selected years (employee, self-employed)	Average, median and deciles for earnings	Average weekly hours	ILO Unemployment rate	Number of vacancies, Hard-to-fill vacancies, Skills shortage vacancies, Occupation	Skills, Abilities, Interests	Structure and descriptions of occupations	(Available through fuzzy search)
Dimensions*	Occupation, Industry, Qualification, Geography, Gender, Status	Occupation, Industry, Qualification, Geography, Gender, Status	Occupation, Industry, Qualification, Geography, Gender, Status	Occupation, Industry, Qualification, Geography, Gender, Status, Age	Occupation, Industry, Geography, Gender, Status	Occupation, Industry, Qualification, Geography, Gender, Status	Occupation, Industry, Geography	Occupation	Occupation	Occupation
Period	2000-2012 #	2012-2022 #	2000-2022 #	2012-2013	2012-2013	2011, 2012, 2013	2011, 2013	2014	2010	2014
Available updates (if known)	every 2-3 years	every 2-3 years	every 2-3 years	annually (or quarterly if required)	annually	annually (or quarterly if required)	every 2-3 years	every 2-3 years	only required when SOC is updated	constant

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 2000-2011 SOC2010 2-digit data are only available

Institute for Employment Research, University of Warwick, *Version 19 June 2014*

1. *Working Futures* – Employment (historical and projected) and Replacement Demands

Description of the dataset and provenance

Historical and projected estimates of employment levels by detailed 4 digit occupational category also covering highest qualification held, industry, region, gender and employment status as published by the UK Commission for Employment and Skills (UKCES).

Details of the owner / curator

Prepared on behalf of the UKCES by the Warwick Institute for Employment Research. The *Working Futures* database draws on official data published by the Office for National Statistics (ONS), including the Business Register and Employment Survey (BRES) and the Labour Force Survey. Details can be found in the Technical Report at:

https://www.gov.uk/government/publications/uk-labour-market-projections-2010-to-2020

https://www.gov.uk/government/publications/working-futures-2012-to-2022

Known quality issues with data

The *Working Futures* employment database provides the most detailed and consistent picture of employment structure available in the UK. It covers the period 1980-2022. It is based on a combination of official sources. The detailed numbers are constructed estimates based on econometric and other techniques rather than simple survey estimates.

Values for 2012-2022 are projections based on a detailed macroeconomic forecast and a set of assumptions about employment prospects as set out in the report at:

https://www.gov.uk/government/publications/working-futures-2012-to-2022

Quality control processes

The employment results in *Working Futures* are subject to a detailed peer review process by UKCES and other stakeholders and users. The results have previously been made freely available to users via the UKCES but subject to obtaining a Chancellor of the Exchequer's Notice to avoid problems of disclosure. The current data set has been aggregated across industries to avoid problems of disclosure

Accuracy of data

Because the employment estimates are a complex combination of information from a number of different sources it is not possible to put precise confidence intervals around the point estimates. Based on guidelines produced by ONS for general use of LFS data (which lie at the heart of the database) the following "rules of thumb" are suggested for users of the data:

Institute for Employment Research, University of Warwick, *Version 19 June 2014*

1. If the numbers employed in a particular category / cell (defined by the 12 regions, gender, status, occupation, qualification and industry (75 categories)) are below 1,000 then a query will return "no reliable data available" and offer to go up a level of aggregation across one or more of the main dimensions (e.g. UK rather than region, some aggregation of industries rather than the 75 level, or SOC 2 digit rather than 4 digit).

2. If the numbers employed in a particular category / cell (defined as in 1.) are between 1,000 and 10,000 then a query will return the number but with a flag to say that this estimate is based on a relatively small sample size and if the user requires more robust estimates they should go up a level of aggregation across one or more of the main dimensions (as in 1).

3. Rounding of estimates - in order to avoid false impressions of precision the API rounds up the estimates before delivering the answer to any query. In the case of the *Working Futures* employment estimates any numbers are rounded to the nearest thousand.

Frequency of update

- Does the data underlying the API change over time?
  - The *Working Futures* series of labour market assessments have been conducted once every 2-3 years since 2002. The next full update is expected to be published in 2017.
  - The underlying data sources (BRES, LFS) are updated on a more frequent basis but individually cannot provide the level of detail available in *Working Futures*.

- Will this data go out of date?
  - The *Working Futures* data are as accurate as they can be at the time they are produced. As time goes by they become more out of date but they are still likely to produce a good broad brush picture of employment opportunities available.

- Does the data you capture change on at least a daily basis?
  - No – see above.

- What type of dataset series is this?
  - Constructed data set based on time series information from various sources including BRES and the LFS

- Is a feed of changes made available?
  - No see above.

- How frequently do you create a new release?
  - When *Working Futures* is updated (every 2-3 years, see above).

- What is the delay between creating a dataset and publishing it?
  - Once the *Working Futures* database has been update it can be uploaded to the LMI for All web portal in a few weeks.

Institute for Employment Research, University of Warwick, *Version 19 June 2014*

- Do you also provide dumps of the dataset?
  - UKCES currently provide access to the main *Working Futures* database in the form of Excel Workbooks. This only goes down to the 2 digit SOC 2010 level. In the future the data may be made available in other ways.

- How frequently do you create a new dump?
  - Infrequently, when the *Working Futures* database is updated.

- Will this data be corrected if it contains errors?
  - yes, subject to resource constraints

Disclosure and confidentiality

UKCES complies with all applicable Data Protection laws in the UK.

The *Working Futures* data included in this tool is non-disclosive.

Institute for Employment Research, University of Warwick, *Version 19 June 2014*

2. *ASHE /LFS* – Earnings and pay

Description of the dataset and provenance

Information on weekly pay (average, median and decile) is taken from a combination of two sources: the Annual Survey of Hours and Earnings (ASHE); and the Labour Force Survey (LFS) (both conducted by the Office for National Statistics (ONS)). ASHE is widely regarded as the most reliable source of information on Pay and Hours, however it does not include information on pay by qualification as well as some other characteristics (such as self employment). This information is available in the LFS.

Details of the owner / curator

Although the ASHE data set is based on a relatively large sample, this is not large enough to produce reliable data at the level of detail ideally required. There are also concerns about information being disclosive. Similar problem apply to the LFS. To avoid these problems the raw survey data are not used. Instead a set of estimates   have been prepared on behalf of the UKCES by the Warwick Institute for Employment Research based on the available information and constrained to match published figures. The estimates are based on published ASHE data and the publically available version of the LFS. However, the estimates presented in the database are predictions from an econometric analysis rather than the raw survey results. This avoids problems of breach of confidentiality and disclosure. The results are constrained to match the published totals using an iterative RAS process. The main predictions are made using a standard "Mincerian" earnings equation.

The characteristics of the groups concerned distinguish: Gender; Industry (75 almost 2 digit 2007 classification categories); Qualifications (highest held, 9 NQF categories); Region/ Country (4 UK countries and 9 English regions); Age (aggregated groups); and Occupation (SOC2010 4 digit categories), for 2012 and 2013.

Known quality issues with data

Both ASHE and the LFS provide robust estimates, but these are subject to sampling errors when sample sizes are small.

Quality control processes

The API suppresses sample cells with zero or small sample sizes.

Accuracy of data

Precise confidence intervals are not provided around the point estimates. Based on guidelines produced by ONS for general use of LFS data the following "rules of thumb" have been adopted:

1. If the numbers *employed* in a particular category / cell (defined by the 12 regions, gender, status, occupation, qualification and industry (75 categories)) are below 1,000 then a query about the related *Weekly Pay* will

return "no reliable data available" and offer to go up a level of aggregation across one or more of the main dimensions (e.g. UK rather than region, some aggregation of industries rather than the 75 level, or SOC 2 digit rather than 4 digit).

2. If the numbers *employed* in a particular category / cell (defined as in 1.) are between 1,000 and 10,000 then a query on the *Weekly Pay* will return the estimated figure but with a flag to say that this is based on a relatively small sample size and if the user requires more robust estimates they should go up a level of aggregation across one or more of the main dimensions (as in 1).

3. Rounding of estimates - in order to avoid false impressions of precision the API rounds up the estimates before delivering the answer to any query. In the case of the *Weekly Pay* estimates they are rounded to the nearest £10.

Frequency of update

- Does the data underlying the API change over time?
    - ASHE is conducted on an annual basis (although the LFS is carried out more frequently). The data could in principle be updated annually (but this would require the processing of the data described above under (Details of the owner / curator)).

- Will this data go out of date?
    - The data are as accurate as they can be at the time they are produced. As time goes by they become more out of date but they can be updated regularly.

- Does the data you capture change on at least a daily basis?
    - No – see above.

- What type of dataset series is this?
    - Time series information based on a cross-sectional survey of employers (ASHE and LFS).

- Is a feed of changes made available?
    - No, see above.

- How frequently do you create a new release?
    - To be decided (annually?)

- What is the delay between creating a dataset and publishing it?
    - To be decided – once the data have been processed they can be uploaded to the LMI for All web portal in a week or so.

- Do you also provide dumps of the dataset?
    - No?

- How frequently do you create a new dump?
    - To be decided – annually?

Institute for Employment Research, University of Warwick, *Version 19 June 2014*

- Will this data be corrected if it contains errors?
  - yes, subject to resource constraints

Disclosure and confidentiality

UKCES complies with all applicable Data Protection laws in the UK.

The *Average Weekly Pay* data included in this tool is non-disclosive.

Institute for Employment Research, University of Warwick, *Version 19 June 2014*

3. *ASHE* – Average Weekly Hours

Description of the dataset and provenance

Information on average weekly hours is taken from the Annual Survey of Hours and Earnings (ASHE) conducted by the Office for National Statistics (ONS). ASHE is the most reliable source of information on Pay and Hours.

Details of the owner / curator

Although the ASHE data set is based on a relatively large sample, this is not large enough to produce reliable data at the level of detail ideally required. There are also concerns about information being disclosive. To avoid these problems the raw survey data are not used. Instead a set of estimates   have been prepared on behalf of the UKCES by the Warwick Institute for Employment Research based on the available information and constrained to match published figures. The estimates are based on published ASHE data but the detailed estimates are predictions based on a simple set of assumptions that differentiates across each of the main dimensions/characteristics. The results are constrained to match the published totals using an iterative RAS process.

The characteristics of the groups concerned distinguish: Gender; Industry (75 almost 2 digit 2007 classification categories); Region/ Country (4 UK countries and 9 English regions); and Occupation (SOC2010 4 digit categories), for 2012 and 2013.

Known quality issues with data

ASHE provides robust estimates, but these are subject to sampling errors when sample sizes are small.

Quality control processes

The API suppresses sample cells with zero or small sample sizes.

Accuracy of data

Precise confidence intervals are not provided around the point estimates. Based on guidelines produced by ONS for general use of LFS data the following "rules of thumb" have been adopted:

1. If the numbers *employed* in a particular category / cell (defined by the 12 regions, gender, status, occupation, qualification and industry (75 categories)) are below 1,000 then a query about the related *average weekly hours* will return "no reliable data available" and offer to go up a level of aggregation across one or more of the main dimensions (e.g. UK rather than region, some aggregation of industries rather than the 75 level, or SOC 2 digit rather than 4 digit).

2. If the numbers *employed* in a particular category / cell (defined as in 1.) are between 1,000 and 10,000 then a query on the *average weekly hours* will

Institute for Employment Research, University of Warwick, *Version 19 June 2014*

return the estimated figure but with a flag to say that this is based on a relatively small sample size and if the user requires more robust estimates they should go up a level of aggregation across one or more of the main dimensions (as in 1).

3. Rounding of estimates - in order to avoid false impressions of precision the API rounds up the estimates before delivering the answer to any query. In the case of the *average weekly hours* estimates they are rounded to the nearest hour.

Frequency of update

- Does the data underlying the API change over time?
  - ASHE is conducted on an annual basis and the data could in principle be updated at that frequency (but this would require the processing of the data described above (under Details of the owner / curator)).

- Will this data go out of date?
  - The data are as accurate as they can be at the time they are produced. As time goes by they become more out of date but they can be updated regularly.

- Does the data you capture change on at least a daily basis?
  - No – see above.

- What type of dataset series is this?
  - Time series information based on a cross-sectional survey of employers (ASHE).

- Is a feed of changes made available?
  - No, see above.

- How frequently do you create a new release?
  - To be decided (annually?)

- What is the delay between creating a dataset and publishing it?
  - To be decided – once the data have been processed they can be uploaded to the LMI for All web portal in a week or so.

- Do you also provide dumps of the dataset?
  - No?

- How frequently do you create a new dump?
  - To be decided – annually?

- Will this data be corrected if it contains errors?
  - yes, subject to resource constraints

Disclosure and confidentiality

UKCES complies with all applicable Data Protection laws in the UK.

The *average weekly hours* data included in this tool is non-disclosive.

4. *Labour Force Survey* – Unemployment Rate

Description of the dataset and provenance

Historical estimates of unemployment rates by detailed 4 digit occupational category, also covering highest qualification held, industry, region, gender and employment status based on the official public version of the Labour Force Survey (LFS) produced and published by the Office for National Statistics (ONS).

Details of the owner / curator

The LFS sample is insufficiently large to produce reliable estimates at the level of detail required. Estimates have been prepared on behalf of the UKCES by the Warwick Institute for Employment Research based on the publically available version of the LFS. Unemployment rates based on the standard ILO definition are computed from the survey variable "estat" which gives the employment status of the individual respondent concerned.

To get the unemployment rates, the number of people in a particular category reporting that they are ILO unemployed from the variable 'estat' are divided by the total number in the labour force (the sum of those reporting "estat" values 1 through 5.

The characteristics of the groups concerned distinguish: Gender; Industry (almost 2 digit 2007 classification categories); Qualifications (highest held, 9 NQF categories); Region/ Country (4 UK countries and 9 English regions); and Occupation (SOC2010 4 digit categories), for 2011, 2012 and 2013.

Information for industry and occupation is based on the individual's previous job. Information on unemployment rates by industry and occupation are based on an individual's previous job.

Data for earlier years are also available but these use different systems for classifying occupations.

Known quality issues with data

The LFS provides robust estimates, but these are subject to sampling errors when sample sizes are small.

Quality control processes

The API suppresses sample cells with zero or small sample sizes.

Accuracy of data

Precise confidence intervals are not provided around the point estimates. Based on guidelines produced by ONS for general use of LFS data the following "rules of thumb" are suggested for users of the data:

Institute for Employment Research, University of Warwick, *Version 19 June 2014*

1. If the numbers *employed* in a particular category / cell (defined by the 12 regions, gender, status, occupation, qualification and industry (75 categories)) are below 1,000 then a query about the related *Unemployment rate* will return "no reliable data available" and offer to go up a level of aggregation across one or more of the main dimensions (e.g. UK rather than region, some aggregation of industries rather than the 75 level, or SOC 2 digit rather than 4 digit).
2. If the numbers *employed* in a particular category / cell (defined as in 1.) are between 1,000 and 10,000 then a query on the *Unemployment rate* will return the estimated rate but with a flag to say that this is based on a relatively small sample size and if the user requires more robust estimates they should go up a level of aggregation across one or more of the main dimensions (as in 1).
3. Rounding of estimates - in order to avoid false impressions of precision the API rounds up the estimates before delivering the answer to any query. In the case of the LFS unemployment rate estimates they are rounded to the nearest percentage point.

Frequency of update

- Does the data underlying the API change over time?
    - The LFS is conducted on a quarterly basis and the data could in principle be updated at that frequency (but this would require the processing of the data described above under (Details of the owner / curator)).

- Will this data go out of date?
    - The data are as accurate as they can be at the time they are produced. As time goes by they become more out of date but they can be updated regularly and frequently.

- Does the data you capture change on at least a daily basis?
    - No – see above.

- What type of dataset series is this?
    - Time series information based on a cross-sectional survey of households (the LFS).

- Is a feed of changes made available?
    - No see above.

- How frequently do you create a new release?
    - To be decided (annually?)

- What is the delay between creating a dataset and publishing it?
    - To be decided - it can be uploaded to the LMI for All web portal in a few weeks.

- Do you also provide dumps of the dataset?

- o No?

- How frequently do you create a new dump?
  - o To be decided – annually / quarterly?

- Will this data be corrected if it contains errors?
  - o yes, subject to resource constraints

Disclosure and confidentiality

UKCES complies with all applicable Data Protection laws in the UK.

The Unemployment rate data included in this tool is non-disclosive.

Institute for Employment Research, University of Warwick, *Version 19 June 2014*

5. *National Employer Skills Survey* – Skill Shortage Vacancies

Description of the dataset and provenance

The UKCES Employer Skills Survey (ESS) collects detailed information on skill deficiencies including skill shortage vacancies with occupations classified using SOC2010. They survey has been conducted in similar form every 2-3 years for over a decade. The survey is intended to produce robust estimates of the total number of vacancies, hard-to-fill vacancies and skill shortage vacancies in the UK from a sample of establishments.

Details of the owner / curator

A special interrogation of the data set to extract 4 digit SOC information has been carried out The most detailed geographical breakdown available is to regions in England and the other nations of the UK: Wales, Scotland and Northern Ireland.

Time period: 2011 and 2013. The UKCES Employer Skills Survey is conducted every two years. The 2011 ESS was the first survey to cover the entire UK.

Known quality issues with data

The survey does not collect data on the numbers employed in each occupation. Therefore the indicators that are possible to generate are limited to the number of vacancies, hard-to-fill and skill shortage vacancies and the percentage of total vacancies which are hard-to-fill and skill shortage.

It also provides information on only a subset of all vacancies.

The overall sample size is relatively large but when it comes down to focusing on detailed occupations the samples are quite limiting.

Quality control processes

The API suppresses sample cells with zero or small sample sizes.

Accuracy of data

Precise confidence intervals are not provided around the point estimates. The following "rules of thumb" have been adopted:

1. If the number of establishments involved in generating and estimate is below 50 then a query about the related *Skill Shortage Vacancies* will return "no reliable data available" and offer to go up a level of aggregation across one or more of the main dimensions (e.g. UK rather than region, some aggregation of industries rather than the 75 level, or SOC 2 digit rather than 4 digit).
2. Rounding of estimates - in order to avoid false impressions of precision the API rounds up the estimates before delivering the answer to any query. In the

Institute for Employment Research, University of Warwick, *Version 19 June 2014*

case of the *Skill Shortage Vacancies* estimates they are rounded to the nearest whole vacancy

Frequency of update

- Does the data underlying the API change over time?
    - ESS is conducted once every 2-3 years. The data could in principle be updated at this frequency annually (but this would require the processing of the data described above under (Details of the owner / curator)).

- Will this data go out of date?
    - The data are as accurate as they can be at the time they are produced. As time goes by they become more out of date, but they can be updated when a new survey is carried out.

- Does the data you capture change on at least a daily basis?
    - No – see above.

- What type of dataset series is this?
    - Irregular time series information based on a series of cross-sectional surveys of employers (ESS).

- Is a feed of changes made available?
    - No, see above.

- How frequently do you create a new release?
    - To be decided.

- What is the delay between creating a dataset and publishing it?
    - To be decided – once the data have been processed they can be uploaded to the LMI for All web portal in a week or so.

- Do you also provide dumps of the dataset?
    - No?

- How frequently do you create a new dump?
    - To be decided

- Will this data be corrected if it contains errors?
    - yes, subject to resource constraints

Disclosure and confidentiality

UKCES complies with all applicable Data Protection laws in the UK.

The *Skill Shortage Vacancies* data included in this tool is non-disclosive

Institute for Employment Research, University of Warwick, *Version 19 June 2014*

6. *O*NET* – Skills, Abilities and Interests

Description of the dataset and provenance

The US have been collecting and developing information on the skills, abilities and interests associated with different jobs for many years. This is all collected together on the US O*NET database, details for which can be found at: http://www.onetcenter.org/

The US data are organised around the detailed US Standard Occupational Classification. By developing a link between this and UK SOC categories, the very rich US information can be exploited. This all assumes that what is relevant for an occupation in the US also applies to the nearest equivalent occupation in the UK.

The LMI for All version of the dataset include three main indicators:

- **Abilities** – These data are ability scores for O*NET SOC codes (occupations). The information shows the level of abilities required and the importance of these abilities for the occupation concerned.
- **Skills** – These data are skills scores for O*NET SOC codes (occupations). The information shows both the levels of skill required and the importance of these skills for the occupation concerned.
- **Interests** – These data show the Interest data associated with each O*NET-SOC occupation.

Defining the indicators: Abilities

Enduring attributes of the individual that influence performance.

For more detail see: http://www.onetonline.org/find/descriptor/browse/Abilities/

- **Cognitive Abilities** (21 elements) — Abilities that influence the acquisition and application of knowledge in problem solving
- **Physical Abilities** (9 elements) — Abilities that influence strength, endurance, flexibility, balance and coordination
- **Psychomotor Abilities** (10 elements) — Abilities that influence the capacity to manipulate and control objects
- **Sensory Abilities** (12 elements) — Abilities that influence visual, auditory and speech perception

Defining the indicators: Skills

For more detail see: http://www.onetonline.org/skills/

Basic Skills –Developed capacities that facilitate learning or the more rapid acquisition of knowledge

Institute for Employment Research, University of Warwick, *Version 19 June 2014*

- Active Learning — Understanding the implications of new information for both current and future problem-solving and decision-making.
- Active Listening — Giving full attention to what other people are saying, taking time to understand the points being made, asking questions as appropriate, and not interrupting at inappropriate times.
- Critical Thinking — Using logic and reasoning to identify the strengths and weaknesses of alternative solutions, conclusions or approaches to problems.
- Learning Strategies — Selecting and using training/instructional methods and procedures appropriate for the situation when learning or teaching new things.
- Mathematics — Using mathematics to solve problems.
- Monitoring — Monitoring/Assessing performance of yourself, other individuals, or organizations to make improvements or take corrective action.
- Reading Comprehension — Understanding written sentences and paragraphs in work related documents.
- Science — Using scientific rules and methods to solve problems.
- Speaking — Talking to others to convey information effectively.
- Writing — Communicating effectively in writing as appropriate for the needs of the audience.

Complex Problem Solving Skills – Developed capacities used to solve novel, ill-defined problems in complex, real-world settings

- Complex Problem Solving — Identifying complex problems and reviewing related information to develop and evaluate options and implement solutions.

Resource Management Skills – Developed capacities used to allocate resources efficiently

- Management of Financial Resources — Determining how money will be spent to get the work done, and accounting for these expenditures.
- Management of Material Resources — Obtaining and seeing to the appropriate use of equipment, facilities, and materials needed to do certain work.
- Management of Personnel Resources — Motivating, developing, and directing people as they work, identifying the best people for the job.
- Time Management — Managing one's own time and the time of others.

Social Skills – Developed capacities used to work with people to achieve goals

- Coordination — Adjusting actions in relation to others' actions.
- Instructing — Teaching others how to do something.
- Negotiation — Bringing others together and trying to reconcile differences.
- Persuasion — Persuading others to change their minds or behavior.
- Service Orientation — Actively looking for ways to help people.
- Social Perceptiveness — Being aware of others' reactions and understanding why they react as they do.

Systems Skills – Developed capacities used to understand, monitor, and improve socio-technical systems

- Judgment and Decision Making — Considering the relative costs and benefits of potential actions to choose the most appropriate one.
- Systems Analysis — Determining how a system should work and how changes in conditions, operations, and the environment will affect outcomes.

- Systems Evaluation — Identifying measures or indicators of system performance and the actions needed to improve or correct performance, relative to the goals of the system.

Technical Skills – Developed capacities used to design, set-up, operate, and correct malfunctions involving application of machines or technological systems

- Equipment Maintenance — Performing routine maintenance on equipment and determining when and what kind of maintenance is needed.
- Equipment Selection — Determining the kind of tools and equipment needed to do a job.
- Installation — Installing equipment, machines, wiring, or programs to meet specifications.
- Operation and Control — Controlling operations of equipment or systems.
- Operation Monitoring — Watching gauges, dials, or other indicators to make sure a machine is working properly.
- Operations Analysis — Analyzing needs and product requirements to create a design.
- Programming — Writing computer programs for various purposes.
- Quality Control Analysis — Conducting tests and inspections of products, services, or processes to evaluate quality or performance.
- Repairing — Repairing machines or systems using the needed tools.
- Technology Design — Generating or adapting equipment and technology to serve user needs.
- Troubleshooting — Determining causes of operating errors and deciding what to do about it.

Defining the indicators: Interests

Preferences for work environments and outcomes.

For more detail see: http://www.onetonline.org/find/descriptor/browse/Interests/

- **Realistic** – Realistic occupations frequently involve work activities that include practical, hands-on problems and solutions. They often deal with plants, animals, and real-world materials like wood, tools, and machinery. Many of the occupations require working outside, and do not involve a lot of paperwork or working closely with others.
- **Investigative** – Investigative occupations frequently involve working with ideas, and require an extensive amount of thinking. These occupations can involve searching for facts and figuring out problems mentally.
- **Artistic** – Artistic occupations frequently involve working with forms, designs and patterns. They often require self-expression and the work can be done without following a clear set of rules.
- **Social** – Social occupations frequently involve working with, communicating with, and teaching people. These occupations often involve helping or providing service to others.
- **Enterprising** – Enterprising occupations frequently involve starting up and carrying out projects. These occupations can involve leading people and making many decisions. Sometimes they require risk taking and often deal with business.

Institute for Employment Research, University of Warwick, *Version 19 June 2014*

- **Conventional** – Conventional occupations frequently involve following set procedures and routines. These occupations can include working with data and details more than with ideas. Usually there is a clear line of authority to follow.

Details of the owner / curator

O*NET is owned and managed by the US Bureau of Labor Statistics (BLS).

Known quality issues with data

The O*NET data are US focussed. However, they have been widely used in many other countries. To the extent that occupations are similar in the UK and the US they will provide useful and relevant information.

Note this is not necessarily a unique mapping – a 4 digit SOC code may link to more than one O*NET occupation. For this reason it is not possible to generate results for more aggregate SOC categories (e.g. 3 or 2 digit).

Quality control processes

The O*NET data are subject to continuous review and validation.

The link between the US and UK occupational systems has been developed using CASCOT a well-established tool used by ONS and various other agencies to classify occupations.in the UK.

Accuracy of data

The key issue with the O*Net data is the relevance of the information to a UK context. It is thought that for most occupations the Skill, Abilities and interests information is relevant.

Frequency of update

- Does the data underlying the API change over time?
  - The *O*NET* data are regularly updated by the BLS every 2-3 years. The next full update is expected to be published in 2016/2017.

- Will this data go out of date?
  - The kinds of information in the database are not expected to change rapidly.

- Does the data you capture change on at least a daily basis?
  - No – see above.

- What type of dataset series is this?
  - Qualitative information on the nature of different types of jobs based on detailed analysis conducted in the US.

- Is a feed of changes made available?
  - No see above.

Institute for Employment Research, University of Warwick, *Version 19 June 2014*

- How frequently do you create a new release?
    - When *O*NET* is updated (every 2-3 years, see above).

- What is the delay between creating a dataset and publishing it?
    - Once the *O*NET* database has been updated it can be uploaded to the LMI for All web portal in a week or so.

- Do you also provide dumps of the dataset?
    - The BLS provides dumps of the data: see: http://www.onetcenter.org/dl_files/DataDictionary15_0.pdf page 17.

    - The LMI for all file was originally downloaded from: http://www.onetcenter.org/download/database?d=db_15_0.zip

- How frequently do you create a new dump?
    - Infrequently, when the *O*NET* database is updated.

- Will this data be corrected if it contains errors?
    - yes, subject to resource constraints

Disclosure and confidentiality

Not applicable.

More information

More details on the datasets can be found in LMI for All - Developing a Careers LMI Database: report (PDF, 1.5 Mb) available at http://www.ukces.org.uk/assets/ukces/docs/supporting-docs/lmi-for-all-phase2a-report.pdf. This report reviews LMI for All up until the end of Phase 2a (Oct 12- March 13) which culminated in the launch of the first release, and makes recommendations for the next stage of development. Also see http://www.ukces.org.uk/ourwork/research/lmi/lmi-for-all#sthash.8qKAg6uc.dpuf

7. *ONS Standard Occupational Classification*

Description of the dataset and provenance

The Standard Occupational Classification (SOC) is a common classification of occupational information for the UK.   Within the context of the classification jobs are classified in terms of their skill level and skill content.   It is used for career information to labour market entrants, job matching by employment agencies and the development of government labour market policies.   SOC2010 is the latest update.

The occupational descriptions dataset provides a detailed structure of SOC2010 occupations together with descriptions for each occupation. An occupational description is an account of the main tasks and duties in a set of jobs, which are characterised by a high degree of similarity. More information is available at: http://www.ons.gov.uk/ons/guide-method/classifications/current-standard-classifications/soc2010/index.html

The following is an example of how the classification is structured:

Major Group	Sub-Major Group	Minor Group	Unit Group	Group Title
(1 digit)	(2 digit)	(3 digit)	(4 digit)	
1	11			**MANAGERS, DIRECTORS AND SENIOR OFFICIALS** **CORPORATE MANAGERS AND DIRECTORS**
		111		**Chief Executives and Senior Officials**
			1115	Chief executives and senior officials
			1116	Elected officers and representatives
		112		**Production Managers and Directors**
			1121	Production managers and directors in manufacturing
			1122	Production managers and directors in construction
			1123	Production managers and directors in mining and energy
		113		**Functional Managers and Directors**
			1131	Financial managers and directors
			1132	Marketing and sales directors
			1133	Purchasing managers and directors
			1134	Advertising and public relations directors
			1135	Human resource managers and directors
			1136	Information technology and telecommunications directors
			1139	Functional managers and directors n.e.c.

More details on the methodology, structure and descriptions are available online:

- Volume 1 (http://www.ons.gov.uk/ons/guide-method/classifications/current-standard-classifications/soc2010/soc2010-volume-1-structure-and-descriptions-of-unit-groups/index.html) outlines the background, resources, concepts, and processes of the Standard Occupational Classification.

Institute for Employment Research, University of Warwick, *Version 19 June 2014*

- Volume 2 (http://www.ons.gov.uk/ons/guide-method/classifications/current-standard-classifications/soc2010/soc2010-volume-2-the-structure-and-index/index.html) consists of a detailed alphabetical index of job titles, giving both the SOC2000 and SOC2010 Unit Group to which each is assigned.
- SOC2010 structure is available at: http://www.ons.gov.uk/ons/guide-method/classifications/current-standard-classifications/soc2010/soc2010-structure--excel-download-.xls

Details of the owner / curator

The Standard Occupational Classification in one of three widely used classifications in the UK used and promoted by the Office for National Statistics (ONS). Other classifications available include the Standard Industrial Classifications (SIC) and National Statistics Socio-economic Classification (NS-SEC). These common statistical frames, definitions and classifications are promoted and used the Office for National Statistics.

See: Elias, P. and Birch, M. (2010). *SOC2010: Revision of the Standard Occupational Classification*, http://www2.warwick.ac.uk/fac/soc/ier/publications/2010/elias_birch_soc2010_revision_2010.pdf

Known quality issues with data

The index may not always yield an appropriate code. Jobs are not static and change with innovation and the introduction of new technologies, changes in the organisation of work, revisions to occupational training and qualification requirements. Therefore, the classification will need to adjusted from time to time to ensure the classification reflects new areas of work, associated training and qualification requirements.

Quality control processes

Not applicable.

Accuracy of data

See above, Known quality issues with data.

Frequency of update

- Does the data underlying the API change over time?
    - Regular reviews of standard classifications are conducted to ensure that economic and social changes are reflected in the classification and, where relevant and possible, that the classification is comparable with European and international standards.
    - To date, a ten year cycle has been adopted by ONS for the revision of the UK national occupational classification.

- Will this data go out of date?
    - The occupational structure and descriptions are as accurate as they can be at the time they are produced. It is reviewed regularly by ONS.

Institute for Employment Research, University of Warwick, *Version 19 June 2014*

- Does the data you capture change on at least a daily basis?
  - Not applicable.

- What type of dataset series is this?
  - Not applicable.

- Is a feed of changes made available?
  - No.

- How frequently do you create a new release?
  - This will be reviewed when a new the standard occupational classification is released.

- What is the delay between creating a dataset and publishing it?
  - To be decided.

- Do you also provide dumps of the dataset?
  - No.

- How frequently do you create a new dump?
  - Not applicable.

- Will this data be corrected if it contains errors?
  - Yes, subject to resource constraints.

Disclosure and confidentiality

UKCES complies with all applicable Data Protection laws in the UK.

Institute for Employment Research, University of Warwick, *Version 19 June 2014*

8. Universal Jobmatch vacancies

Description of the dataset and provenance

Universal Jobmatch is a free service that enables individuals to search for and apply for jobs on one of the largest job boards in Europe. Individuals do not need to be registered to search for jobs but setting up a Universal Jobmatch account will enable them to do much more such as create CVs and complete application forms.

Universal Jobmatch vacancies are continually added to the LMI for All database. These can be accessed using a fuzzy serarch.

Matching vacancies to SOC2010 occupations currently being explored.

Details of the owner / curator

Universal Jobmatch is a service offered through Government Gateway. The Monster Corporation operates the system on behalf of the Department for Work and Pensions (DWP). The website replaced the Jobcentre Plus job search tool and Employer Services Direct.

Known quality issues with data

There are a number of known issues with the data and the search process used to provide results. Some vacancies are known to be false.

Discussions are underway to provide more reliable vacancy data.

Quality control processes

Where errors are found and reported these will be corrected subject to resource constraints. An alternative process of providing vacancy data is underway.

Accuracy of data

See above, Known quality issues with data.

Frequency of update

- Does the data underlying the API change over time?
  - Vacancy data are continuously updated.

- Will this data go out of date?
  - No.

- Does the data you capture change on at least a daily basis?
  - Yes.

- What type of dataset series is this?

Institute for Employment Research, University of Warwick, *Version 19 June 2014*

- o   Not applicable.

- Is a feed of changes made available?
    - o   No.

- How frequently do you create a new release?
    - o   Not applicable.

- What is the delay between creating a dataset and publishing it?
    - o   Not applicable.

- Do you also provide dumps of the dataset?
    - o   No.

- How frequently do you create a new dump?
    - o   Not applicable.

- Will this data be corrected if it contains errors?
    - o   Yes, subject to resource constraints.

Disclosure and confidentiality

UKCES complies with all applicable Data Protection laws in the UK.

Institute for Employment Research, University of Warwick, *Version 19 June 2014*