Please cite this paper as:

Green, T. (2009), "We Need Publishing Standards for Datasets and Data Tables", *OECD Publishing White Papers*, OECD Publishing.
doi:10.1787/787355886123
http://dx.doi.org/10.1787/787355886123

OECD Publishing White Papers

We Need Publishing Standards for Datasets and Data Tables

Toby Green, OECD Publishing, Paris
toby.green@oecd.org

REVISED VERSION FEBRUARY 2010

Acknowledgements

The author would like to thank his colleagues at OECD who have worked hard to develop these proposed standards: Pascale Cissokho-Mutter, Terri Mitton, Eileen Capponi, Matt Brosius, Jerome Cukier Maria Bjurström and Deborah Scott-Douglas.

NOTES ON THE REVISION OF THIS PAPER

An earlier version of this white paper was released under the same title on 20 April 2009. This revised version supersedes the earlier one. Following the release of the first version of this paper in April 2009 (doi: 10.1787/603233448430), it has become clear that the issue of how datasets should be published and linked to scholarly literature has become a widely discussed topic among librarians, data producers and publishers. Since June 2009, the metadata standards proposed in this paper are being put into practice on OECD's new publishing platform, OECD iLibrary (www.oecdiLibrary.org) . This paper has been updated to reflect the bibliographic metadata standards as implemented at OECD and recent developments in the scholarly industry towards establishing standards for datasets and data tables.

Publishing Standards for Datasets and Data Tables

Go to Google.com and search for "journals". The first page of results is full of references to professional publishing websites such as OUP (Oxford University Press), PubMed, AMS (American Mathematical Society) *et al.* Just what one would expect.

Go to Google.com and search for "datasets". On the first page of results you'll find only two references to anything that might be professional – a link to the US Census website and a link to some data posted by researchers at the World Bank. (Interestingly, this World Bank site is not their famous World Development Indicators website.) The other results are a mishmash of poorly presented and poorly maintained pages from universities and other research bodies. None of these sites is presented in a professional manner and each has its share of broken links.

Research is all about gathering data. Academic papers, journal articles and monographs cannot be written without data. Before the Internet, data could not be made available easily, but now datasets are being posted on departmental websites in universities and research centres around the world. But can you find them? Will they be there tomorrow? Judging from the Google search, the answer is not positive.

Despite not appearing in the Google search results, publishers are getting involved. According to a recent ALPSP (Association of Learned and Professional Society Publishers) report on scholarly publishing practice (Cox and Cox, 2008), 45% of journal publishers said they provided access to datasets associated with the journal articles they publish. Since 160 publishers replied to this question, this means that at least 72 journal publishers report they are handling data.

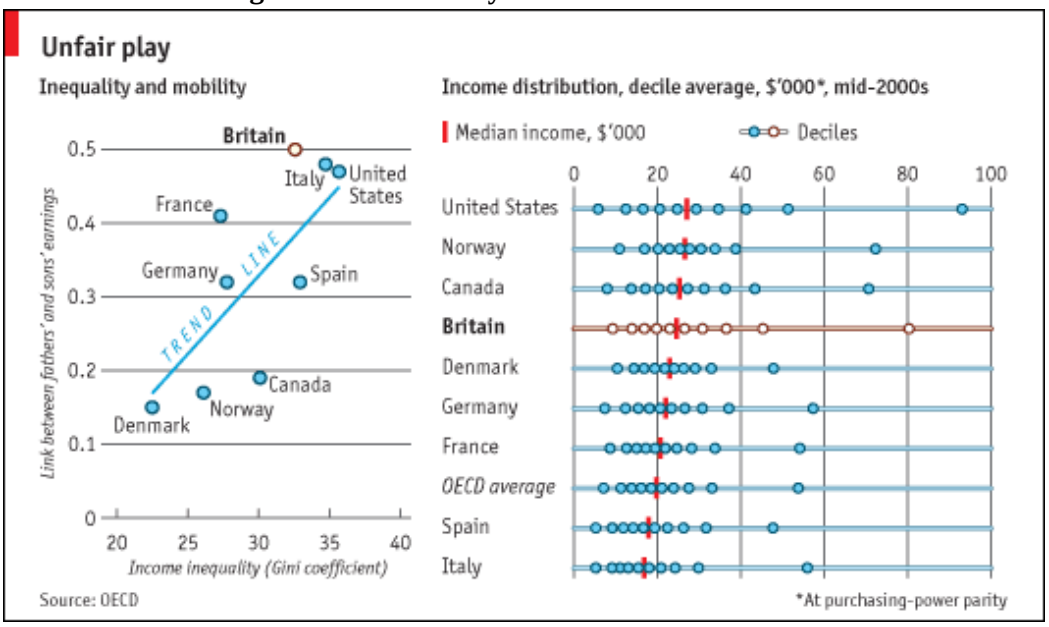
However, dig a little deeper and it's easy to see there are no rules about how to publish, present, cite or otherwise catalogue datasets. Consider Figure 1, published by The Economist. Published as part of a piece called "Pain all round: The importance of fairness in an economic downturn", it gives the reader no clue beyond "Source: OECD" as to which of the 369 datasets OECD publishes it came from.

A search of ScienceDirect shows that many authors are using OECD data in their research. Yet when they need to cite OECD data, they cite the data in a wide variety of ways. They point to:

- OECD print editions (which often have extensive statistical annexes) rather than the original dataset
- the OECD main website (www.oecd.org)
- nothing more than "OECD"
- pages deep in the OECD website, none of which use persistent links

Figure 2 is from a recent issue of Elsevier's journal *World Development*.

Figure 1. A commonly found citation for OECD data



Source: The Economist (2008) "Pain all round, please: The importance of fairness in an economic downturn", *The Economist* (print edition), 23 October 2008.

Figure 2. A commonly found citation for OECD data

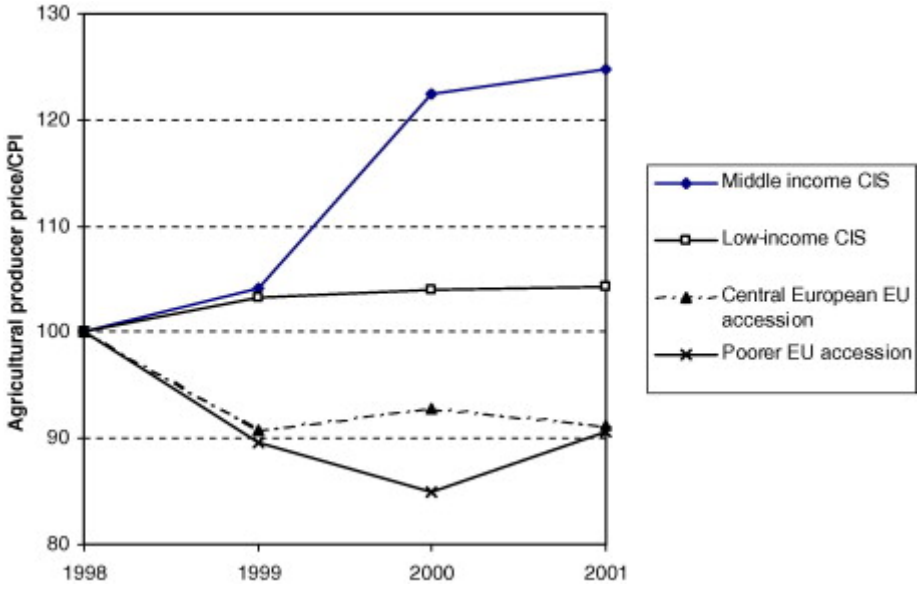


Figure 3. Change of relative agricultural producer prices since 1998. Middle-income CIS show average for Russia, Kazakhstan, and Ukraine. Low-income CIS show average for Azerbaijan, Kyrgyzstan, and Moldova. Central European EU accession countries show average for Czech Republic, Estonia, Hungary, Poland, Slovakia, and Slovenia. Poorer EU accession countries show average for Bulgaria, Latvia, and Lithuania. Source: OECD, 2004 and CIS Statistics, 2003.

Source: Macours, K. and J.F.M. Swinnen (2008), "Rural-Urban Poverty Differences in Transition Countries", *World Development,* Vol. 36, Issue 11, November 2008. doi:10.1016/j.worlddev.2007.11.0043

In another example, an author publishing in the journal *World Development* went to great trouble listing all the data sources in an extensive Appendix. However, some of the items listed would not be of much help to a reader:

> *Main mortality estimate: Estimated settler mortality. Settler mortality is calculated from the mortality rates of European-born soldiers, sailors, and bishops when stationed in colonies. It measures the effects of local diseases on people without inherited or acquired immunities.* **Source*: Acemoglu* et al. *(2001), based on Curtin (1989) and other sources***.
>
> *Tertiary school enrolment: School enrolment, tertiary (% of gross).* **Source*: Barro and Lee (2000) and their databases.***

Source: Lee, K. and B-Y. Kim (2009), "Both Institutions and Policies Matter but Differently for Different Income Groups of Countries: Determinants of Long-Run Economic Growth Revisited", *World Development*, Vol. 37, Issue 3, March 2009. doi:10.1016/j.worlddev.2008.07.004

This is not to single out Elsevier. A similar result is found in RePEc[1], the world's largest collection of papers in economics and on other publishers' e-journals sites. ESDS International[2], a UK data aggregation service, asks its users to cite OECD data as follows:

> *Citation information*
> *Publications based on ESDS data collections should always acknowledge the data source with a citation in the reference section. The bibliographic citation for this database is: Organisation for Economic Cooperation and Development, <Dataset name>, ESDS International, University of Manchester*

Not only is there no linking URL provided, the same data now risk being cited in two different ways depending on where the user sourced the data: one way via the services of OECD, another via those of ESDS.

In view of the different advice received, it is no surprise authors and publishers are so unsure about how they should cite data sources. But what about librarians – are they doing a better job managing and curating datasets?

Librarians take the trouble to catalogue each and every book and journal they acquire – thousands of titles a year. Yet when it comes to datasets the situation looks very messy. Leaving aside confusion about the meaning of the term "dataset" (some librarians list Ingenta Journals, Cambridge Scientific Abstracts, and Inspec as "datasets"), a quick trawl through some library OPACs shows that none have managed to catalogue all 369 datasets published by OECD. In most cases, OECD datasets are bundled together under the brand name we use for our online publishing platform, SourceOECD. In some cases the link is to the OECD Statistics Department website, not the publishing platform where the actual datasets are housed. Again, this is not unique to OECD. Other datasets are either left uncatalogued (how many librarians have

1 www.repec.org

2 www.esds.ac.uk/international/introduction.asp

catalogued the "Delve datasets" from University of Toronto[3]?) or are bundled together in unstructured lists in some corner of a library website. Since datasets are so difficult to catalogue, what chance is there for federated search tools to discover datasets alongside articles and book chapters?

None of this is to criticise authors, publishers, aggregators or librarians. They can hardly do better in the absence of an accepted system for how datasets should be cited and catalogued – but now that datasets are becoming widely available and so many publishers are beginning to get involved (if only to provide links from their journal articles), there is a need for a bibliographic system to help authors cite datasets and for librarians to catalogue them.

In today's world, does this matter? Surely everyone will find what they need from a general search like Google?

But do they? In Inger and Gardner's white paper on how readers navigate to scholarly content (2008) they found that:

- When a reader already has a reference or citation and wishes to read the article on line, a general web search engine ranks fourth behind specialist bibliographic search engines, library web pages and journal homepages as the starting point.
- When keeping up to date with the latest issues of journals, email alerts prove to be the most used, followed by visiting journal homepages and library websites ahead of general search engines.
- When searching for unknown articles on a specific subject, they prefer to search specialist bibliographic databases to general search engines (like Google) and rank library web pages as highly as journal gateways as starting points.

Clearly, if data providers rely on a post-it-and-Google-will-find-it approach, they will miss out on a great deal of traffic from readers who are using alternative routes to discover content.

As e-journal publishers know, and as e-book publishers are finding out, readers want everything to be connected. Readers don't want to find e-books in different e-silos from e-journals. They want all scholarly content, including the underlying datasets, to be interconnected via bibliographies and reference listings.

OECD is not the first to recognise this. Altman and King (2007) proposed a citation standard for scholarly data in 2007 but they did not include cataloguing metadata in their proposal.

In the UK, librarians are getting interested. UK Research Data Service[4] recently completed a feasibility study which recommended a national data management service. Research Libraries UK and the IT directors from the Russell Group of Universities received funding from HEFCE (Higher Education Funding Council for England) and JISC (Joint Information Systems Committee) to undertake the study. As Jean Sykes, librarian and director of IT services at London School of Economics noted, "The data underlying a research project can be extremely important, and can be regarded as part of the research

3 http://www.cs.utoronto.ca/~delve/data/datasets.html

4 http://ukrds.ac.uk

output just as much as the finished experiment or publishing article. Research data ... is an invisible asset...". (Caldwell, 2008) In Australia, the implementation of the Australian National Data Service[5] is already underway.

Proof that readers want to access underlying data

OECD launched a service called StatLinks in 2004. The concept is simple. Under each table, chart and graph appearing in an article or book chapter, a DOI (Digital Object Identifier) link is printed alongside the traditional 'Source' legend. By following the DOI link, readers are able to download a spreadsheet containing the data used to create the table, chart or graph. By 2008, OECD had put 20 000 StatLinks into its publications and in 2008 alone, 980 381 spreadsheet files were downloaded. Proof, if it were needed, that readers do take the chance to get hold of original data when it's offered.

Researchers themselves understand the need to share their data. Aside from the World Bank researchers noted at the beginning of this paper, scholars at Oxford University were polled in the summer of 2008 for their views. The survey showed that informal mechanisms for sharing data were in place, but they were patchy and uneven across disciplines. A similar survey at Bristol University showed the same need for better management of data repositories and decisions on what needs to be stored and how it will be indexed. (Caldwell, 2008)

In Germany, to enable citations and improve retrievability of data, the German Research Foundation (DFG) created a project on Publication and Citation of Scientific Primary Data. Starting with the field of earth science, the German National Library of Science and Technology (TIB) has now established itself as a Digital Object Identification (DOI) registration agency for scientific primary data and claims to have registered 600 000 datasets[6].

TIB has gathered a group of leading research libraries and technical information providers to establish a partnership to improve access to research data on the Internet. The German National Library of Science and Technology (TIB), the British Library, the Library of the ETH Zurich, the French Institute for Scientific and Technical Information (INIST), the Technical Information Centre of Denmark and the Dutch TU (Dutch University of Technology) Delft Library all signed a Memorandum of Understanding (MOU) to this effect during the meeting of the International Council for Scientific and Technical Information (ICSTI) in Paris on 2 March 2009. Recently, the Canadian Institute for Scientific and Technical Information (CISTI) and the Australian National Data Service (ANDS) have also joined as partners, now expanding the group outside of Europe. Their goal is to establish a not-for-profit agency that enables organisations to register research datasets and assign persistent identifiers so that research datasets can be

5 http://ands.org.au

6 http://www.tib-hannover.de/en/special-collections/research-data/

handled as independent, citable, unique scientific objects[7]. The partners officially founded "DataCite" on 1 December 2009 in London[8].

In the United States, to improve the long-term preservation and accessibility of research data, the Dataverse Network Project[9] is being developed at Harvard University by the Institute for Quantititative Science (IQSS). The DataVerse Network provides a platform for preserving, sharing and citing research data using persistent identifiers in place of URLs called universal numerical fingerprints (UNFs). Several institutions, including IQSS, maintain permanent data archives and infrastructures for data.

So, it seems the time is right to propose a bibliographic standard for datasets.

OECD's vision is to make data outputs as accessible and easy to find and use as written outputs like working papers, journal articles and book chapters. Moreover, the vision is to make all published outputs compatible with and discoverable from all scholarly publishing and discovery systems. This means it should be easy for an author writing a paper for a journal to cite a database accurately, in such a way that the publisher can offer a reader a persistent link to the database without having to do anything more than they're doing today to link to articles. Equally, a librarian should be able to catalogue datasets with no more effort than it takes to catalogue and maintain records for e-serials. Readers who like to discover content via specialist bibliographic databases or other specialist websites should also expect to find references to datasets among references for articles, working papers, books and so on. In short, datasets must be as discoverable and as linkable as any scholarly publication.

Inger and Gardner (2008) studied eleven different discovery channels for e-journal articles ranging from the informal (listing on author's departmental page) to the formal (specialist bibliographic databases). Since each of the eleven scored in terms of being helpful to readers in discovering journal articles, to achieve the vision outlined above, a parallel system for data needs to be considered which should be compatible with that used for journals and books.

As Table 1 shows, if bibliographic metadata is prepared to the same standard, most of the discovery channels for e-journals can easily be exploited by datasets. In fact, OECD has been exploiting some of these channels for its datasets since 2001 – often without the channel owner realising what is going on. An example is journal gateways: users of IngentaConnect and SwetsWise can discover OECD datasets among the journal articles. This was achieved by simply producing metadata for each dataset to the same style and standard as used by journals and sending it to these gateways along with OECD journals metadata. All this was done without Ingenta or Swets realising because the metadata structure was the same as that used for journals.

7 http://www.tib-hannover.de/en/the-tib/news/news/id/114/ accessed on 25 March 2009

8 http://www.datacite.org accessed on 17 September 2009

9 http://thedata.org/ accessed on 27 January 2010

Table 1. Comparison of Inger and Gardner's list of discovery channels for e-journals and OECD proposed discovery channels for datasets

Inger and Gardner's list of discovery channels for e-journals	Proposed equivalents for datasets
Specialist bibliographic database	The same – requires bibliographic metadata for datasets to be provided to these databases in a compatible format (*e.g.* Onix)
Library web pages	The same – requires bibliographic metadata for datasets to be provided to these databases in a compatible format (*e.g.* MARC 21, Dublin Core)
Specialist, subject-specific portals	The same – requires bibliographic metadata for datasets to be provided in a compatible format
Key research group website	The same – unlikely to use a bibliographic standard, so simply requires persistent links to the dataset or a widget to be posted
Department webpage on institutional website	The same – unlikely to use a bibliographic standard, so simply requires persistent links to the dataset or a widget to be posted
Publisher's website	The same – if a formal catalogue of other publications is presented, requires bibliographic metadata for datasets in a compatible format; otherwise persistent links to the dataset or a widget to be posted
Email-based alerts	The same – requires bibliographic metadata for datasets with a suitable date-driven trigger when updates are made available
Journal homepage	Database homepage would be the equivalent. This page needs to have its own persistent URL.
Journal gateway[10]	The same – requires bibliographic metadata for datasets to be provided to these gateways in a compatible format
General web search (*e.g.* Google, Yahoo!)	The same – higher rankings will result if bibliographic metadata for datasets is exposed to search engines
Scholarly society website	The same (if appropriate) – if a formal catalogue of other publications is presented, requires bibliographic metadata for datasets in a compatible format; otherwise persistent links to the dataset or a widget to be posted

[10] Examples are IngentaConnect, SwetsWise, EBCSO Host *et al.*

Looking beyond discovery, what other requirements are there for a bibliographic standard for datasets?

Cross-referencing. Launched in 2000, CrossRef now has more than 38.1 million links to scholarly materials registered from 2867 publishers[11]. The bulk of this content is scholarly articles from journals, but CrossRef's mission is to provide the citation linking backbone for all scholarly information in electronic form. This clearly encompasses datasets. The benefits are straightforward. By using CrossRef, dataset publishers can interweave datasets into the scholarly information network alongside journal articles and book chapters. Authors will be able to cite datasets they've used in their manuscripts, confident that their publisher will be able to not only find the link, but maintain the link in the future. Readers will be able to click on references to cited datasets, not just to cited articles, with confidence. Forward linking, now common in journal publishing, would be possible, so users could find published research articles and book chapters which draw on a particular dataset.

Library systems. Books and journals have been catalogued by librarians for generations and international standards have emerged to improve the process. Library catalogues are now on line, and as Inger and Gardner (2008) have shown, they are still used intensively as discovery tools by readers. The main bibliographic standard for library systems is MARC (MAchine Readable Cataloguing). This is a format standard for the storage and exchange of bibliographic records and related information in machine-readable form. In recent years, ONIX has emerged as the international standard for representing and communicating book industry product information in electronic form. ONIX for serials also exists. The key message here is that if standards can help streamline the cataloguing of books and journals, why can't they be extended to include datasets too?

Taking these requirements into account, OECD is proposing to implement a metadata standard for publishing datasets, collections of datasets and individual tables. The detailed proposal can be found in an annex to this paper, but a summary can be seen in Table 2 – and it also shows how OECD's proposal compares with Altman and King's 2007 proposal.

If these fields are stored in an XML or relational database, exporting them to discovery channels and library systems in standard file formats such as ONIX or MARC 21 is a simple procedure. Many of the fields are not needed for citation purposes or for cataloguing but they would be invaluable for users when they come to the dataset's homepage.

Similarly, it is a simple step to create downloadable citations for use in the common bibliographic management systems used by authors, *e.g.* RefWorks and EndNotes.

[11] Source: www.crossref.org on 16 September 2009

Table 2. Comparison of Altman and King's and OECD's proposal for dataset publishing metadata fields

Proposed dataset publishing metadata fields	OECD proposal	Included in Altman and King's proposal?
Unique, persistent, global identifier *e.g.* DOI *(digital object identifier of the content)*	**Mandatory**	Yes
Main title *(in all available languages)*	**Mandatory**	Yes
Subtitle *(in all available languages)*	**Optional**	No
Author(s) **-** *first name* *- last name* *- affiliation*	**Mandatory**	Yes
Publication date *DD-MM-YYYY*	**Mandatory**	Yes
Next publication date (3 fields) *- Day (when available), Month (when available),Year* *to show when the dataset is next due to be updated*	**Mandatory**	No
Periodicity *(e.g. Annual, Monthly, Quarterly, etc.)* *Some datasets have a regular update frequency – shown here*	**Mandatory**	No
Languages of the content (ISO Codes) *Useful when publishing datasets in more than one language*	**Optional**	No
Size number of cells (algorithm to calculate size must be defined) *Useful to show total file size*	**Optional**	No
Countries covered *When the dataset is country specific, this relates to the Country ISO Code of the related country*	**Optional**	No
Period covered: Start year	**Optional**	No
Period covered: End year	**Optional**	No
Variable index - A classification with variable titles	**Mandatory**	No
Short abstract	**Mandatory**	No
Long abstract	**Mandatory**	No
Keyword(s)	**Optional**	No

Table 2. Comparison of Altman and King's and OECD's proposal for dataset publishing metadata fields *(continued)*

Classification *For example JEL classifications in the case of economics*	**Optional**	No
Belongs to (parent-child relationship) *Links a dataset to any other parent dataset to which it is associated. This is useful when a dataset is released which is a sub-set of a larger dataset.*	**Optional**	No
Has main parent *For a dataset belonging to more than one larger dataset, links the dataset to its main parent dataset (as opposed to other "step" parent datasets)*	**Optional**	No
Has physical form Link to dataset in various output formats (MS Excel®, PDF…) *Allows a dataset to be released in different file formats*	**Mandatory**	No
Is related to *Relates complementary products to facilitate linking*	**Optional**	No
External links *Relates objects to external web pages,* e.g. *author's website*	**Optional**	No
Supersedes *Indicates that a dataset has replaced an earlier version*	**Optional**	No
Is continued by *Indicates that a dataset has been discontinued and is replaced by a new one*	**Optional**	No
Is imprinted by *Indicates the legal /organisational body which owns the dataset at imprint level*	**Optional**	No
Is copyrighted by *Indicates the legal /organisational body which owns the publication at copyright level*	**Mandatory**	No
Universal numeric fingerprint *Method for showing if a dataset has changed in any meaningful way since it was initially released. In OECD's case this would be managed by the Supersedes/Continued by fields.*	**Not in OECD's Proposal**	Yes

OECD is proposing to cite datasets in the following way:

· <author(s)> (<year of publication date>), "<main parent sub-collection main title>", <dataset main title >: <dataset subtitle>, <main parent top dataset collection main title> (database).
<doi>
<doi link>
(Accessed on <date>)

Which would give:

· OECD (2009), "Key short-term indicators", *Main Economic Indicators* (database).
doi: 10.1787/data-00039-en
http://dx.doi.org/10.1787/data-00039-en
(Accessed on 14 September 2009)

In addition to datasets, OECD is proposing standards to cite tables. This is important since so many tables, rather than the original datasets, are used as source data by authors. OECD is proposing the following ways to cite tables:

When the table comes from a publication (*e.g.* book):

· <author> (<year of publication date>), "<table number>. <table title>:<table subtitle>" , in <publication title>: <publication subtitle>, <publisher>
<doi>
<doi link>

Which would give:

· Smith, J. (2008), "Figure 1.2. Broadband penetration in OECD countries", in *OECD Communications Outlook 2008*, OECD Publishing
doi: 10.1787/000530172303
http://dx.doi.org/10.1787/000530172303

When the table comes from a stand-alone series of tables rather than a publication:

· <author> <year of publication date>, "<table title>:<table subtitle>", <*table collection title*>, No. <table order number>
<doi>
· <doi link>
· (Accessed on <date>)

Which would give:

· OECD (2009), "Income tax plus employee social security contributions", *OECD Key Tables on Taxation*, No.1.
doi: 10.1787/16097319-2009-1-table1
http://dx.doi.org/10.1787/16097319-2009-1-table1
(Accessed on 21 January 2009)

Dynamic datasets

Many datasets are being updated on a rolling basis, adding new data as and when received. Occasionally, revisions are made to the entire dataset which changes the old data. All of these changes are noted and explained in the statistical metadata (*i.e.* the metadata which describes the data itself, rather than the publishing metadata which is used to describe the dataset). A citation, however, is supposed to link a reader back to the *same* publication which the citing author read. In the case of a dynamic dataset, linking back to the dataset as it was when an author used it to write a paper is clearly impossible. This poses a significant challenge.

The case of dynamic datasets is not the same as versioning. With versioning it is possible to track back to earlier versions as is done with websites like Wikipedia. In the case of dynamic datasets the volume of changes can be so large or frequent to make tracking back impossible to manage.

OECD has discussed this issue with CrossRef and there is no immediate solution. Further discussions in the industry are needed. In the meantime, OECD will use a unique DOI to link to each dataset's homepage, dynamic or not, and use the publishing metadata to alert users to the dynamic nature of the dataset. Details of changes to the data will be found in the associated statistical metadata.

Static tables

OECD already publishes a large number of tables as part of its publications. In 2009 OECD launched a collection of key tables series, for example, a series of key tables on employment and labour markets. All these tables share a common characteristic: they are static objects. This makes them as easily citable as a journal article. However, many tables are updated on a monthly, quarterly or annual basis. This means they are "serials", just like an annual publication. OECD is therefore creating "serial" metadata for these "series" or "collections" so they can be managed and cited just like an annual reference work. This means authors can cite a particular table "edition" (*e.g.* the 2008 edition) and the DOI link will take readers to that "edition". A link on that edition's homepage offers the reader the option of clicking "forward" to the latest, 2010, edition. Backlinks to previous editions will also be provided.

Renditions

Data objects appear in a variety of versions, or renditions corresponding to the different languages of the same original data content and/or to the different types of proposed update methods. Data objects can be updated on a continuous basis where previously released data may be over-written (dynamic object), or can be updated periodically, where each update is a static, unchanging object. For example, a table may be available in two languages (English and French) as two static objects, and for each language be available in HTML, MS Excel® and PDF formats. OECD is proposing to use a single DOI link for each table, linking to the homepage for the data object rendition (*e.g.* the object in English), from which the reader can choose which electronic format to download. The same citation information will be embedded in each rendition. An example of different renditions for individual tables can be seen in OECD's Key Tables service[12].

Implementation of the proposed metadata schema at OECD

In 2009, the Organisation's publishing catalogue was enhanced to manage the bibliographic metadata for datasets and tables alongside other publications. Statistical editors added a layer of publishing metadata and catalogued 369 datasets, 90 key tables and over 9 000 static tables and graphs in analytical books. The proposed metadata schema is being used in OECD's new iLibrary platform, launched in mid-2009. Users can now search and click directly to tables, journals, working papers, books or datasets. OECD aims to add

[12] http://www.oecd.org/statistics/keytables accessed on September 15th 2009.

links to source datasets in publications to interweave statistical datasets and tables into OECD's analytical work.

Users of the OECD iLibrary[13] can easily download citations for datasets and tables in a form compatible with popular bibliographic management systems, including EndNotes, Ref Manager, Pro Cite and BibTex. All of the datasets and tables have DOIs deposited with CrossRef registration agency so that they can be referenced by users and publishers - a critical step to ensure that users will find the data in the future. Every published statistics content in a given language has its own home page, title, DOI, abstract, publication date, periodicity, next update, etc., with links to download or view physical items in a given format/media, and links to related publications. Librarians will be offered standard MARC 21 records for datasets, alongside records for books and periodicals. The publishing metadata will be exported to multiple channels in standard bibliographic formats to other sites, such as RePEc, to online bookshops and librarian catalogues. Users will be able to discover, cite and link to datasets and tables as easily as any other published output of the Organisation.

Conclusion

Datasets are a significant part of the scholarly record and are being published more and more frequently, either formally or informally. Many publishers are beginning to link to them from their journals and authors are trying to cite them in their articles. Librarians would like a way to manage them alongside other publications. In short, they need to be integrated into the scholarly information system so that authors, readers and librarians can use, find and manage them as easily as they do working papers, journal articles and books.

In this paper, OECD is proposing some standards for citing and bibliographic management of datasets and data tables. OECD has built a new online publishing platform, the OECD iLibrary, which hosts working papers, journals, books, tables and datasets. Launched in mid-2009, this platform uses the standards proposed above. Librarians are offered MARC 21 records for datasets, alongside records for OECD books and periodicals. Users of the platform are invited to download citations for datasets and tables in a form compatible with popular bibliographic management systems. All the DOIs for the datasets and tables have been deposited with CrossRef registration agency, for other publishers to use.

Following the release of the first version of this paper in April 2009, it has become clear that the issue of how datasets should be published and linked to scholarly literature has become a widely discussed topic among librarians, data producers and publishers. The hope expressed in the earlier version of this paper that the "scholarly information industry accepts that datasets are a vital part of the scholarly record" is no longer a hope. It has become an objective as evidenced by the creation of DataCite and the news that NISO (National Information Standards Association) is preparing a proposal for a US National Science Foundation grant to fund a project to look at both metadata standards and long-term archiving. This is very welcome and OECD looks forward to working with both groups, and others, to develop a metadata standard for datasets.

[13] www.oecdilibrary.org

Bibliography

Altman, M. and G. King (2007), "A Proposed Standard for the Scholarly Citation of Quantitative Data", *D-lib Magazine*, March/April 2007, Vol. 13 No.3/4

Caldwell, T. (2008), "Academia's Buried Treasure", *Information World Review*, 8 October 2008

Cox, J. and L. Cox (2008), "Scholarly Publishing Practice 3: Academic Journal Publishers' Policies and Practices in Online Publishing, 3rd Edition", ALPSP, London

Inger, S. and T. Gardner (2008), "How Readers Navigate to Scholarly Content", http://www.sic.ox14.com/howreadersnavigatetoscholarlycontent.pdf

Lee, K. and B-Y. Kim (2009), "Both Institutions and Policies Matter but Differently for Different Income Groups of Countries: Determinants of Long-Run Economic Growth Revisited", *World Development*, Vol. 37, Issue 3, March 2009. doi:10.1016/j.worlddev.2008.07.004

Macours, K. and J.F.M. Swinnen (2008), "Rural-Urban Poverty Differences in Transition Countries", *World Development*, Vol. 36, Issue 11, November 2008. doi:10.1016/j.worlddev.2007.11.0043

Annex 1. Dataset metadata

Definition of dataset

A content type (consistent set of related data such as a multi-dimensional data cube) published:
* as part of a collection or
* stand-alone (in this case it can be subject to subscription and managed as a serial)
has a DOI (DOI suffix = data**-<Number on 5 digits>-ISO code language** *e.g.* data-00002-en, data-00023-fr)

A dataset can belong to more than one collection of datasets. One collection of datasets is necessarily the **default (*i.e.* "main")** parent while the eventual others should be step-parents.
The **default** parent is always attached to the dataset when the dataset is accessed straightforwardly on line (search, DOI, etc.). The step-parent is displayed only when the dataset is accessed from the Table of Contents of the step-parent > see link **Is Default Parent Of** at level of database component.

Main metadata exist at 3 levels: conceptual, rendition, item.

- This object is cited.
- This object is searchable on the OECD iLibrary

Datasets belonging to statistical collection are cited as follows:
<author(s)> (<year of publication date>), "<dataset main title>: <dataset subtitle>", <*default parent collection main title*> (database).
doi: <doiprefix>/<doisuffix>
<doi URL>
(Accessed on <date>)
where (database) is a label to be displayed
where date is formatted as follows: dd month label yyyy

Which would give:
OECD (2008), "Social expenditure aggregates", *OECD Social Expenditure Statistics* (database).
doi: 10.1787/data-00001
http://dx.doi.org/10.1787/data-00001
(Accessed on 21 December 2008)

Which would give (in this example there is no dataset subtitle and a joint OECD and FAO copyright):
OECD/FAO (2008), "World prices", A*gricultural Outlook* (database).
doi: 10.1787/data-00002
http://dx.doi.org/10.1787/data-00002
(Accessed on 21 December 2008)

Stand-alone datasets are cited as follows:
<author(s)> (<year of publication date>), <*dataset main title*>: <*dataset subtitle*>(database).
doi: <doiprefix>/<doisuffix>.
<doi URL>
(Accessed on <date>)

Which would give (in this example there is no dataset subtitle and a single copyright owner):
OECD (2008), *OECD Telecommunications Statistics* (database).
doi: 10.1787/data-00001.
http://dx.doi.org/10. 1787/data-00001.
(Accessed on 21 December 2008)

Dataset metadata	Status
ISSN ***Relevant only for stand-alone datasets which are subject to subscription***	**Optional**
Dataset code *(unique identifier of the dataset in its host location)*	**Optional**
DOI *(digital object identifier of the content)* Each dataset in one language should have its own DOI. - The DOI of the English dataset should resolve to the library homepage of the dataset on the English interface. - The DOI of the equivalent French dataset should resolve to the library homepage of the dataset on the French interface.	**Mandatory**
DOI number (required for management of DOI suffix syntax)	**Mandatory**
Main title *(in all available languages)* *Displayed on the iLibrary in citation, search results, homepages*	**Mandatory**
Subtitle *(in all available languages)*	**Optional**
Author(s) ***** Institutional author will be OECD unless external. ***** Physical author **-** *first name* *- last name* *- affiliation* *- order (when several authors - managed by OECD)* A logo in different sizes may be associated to institutional authors.	**Optional**
Author(s) - order (when several authors - managed by OECD)	**Optional**
Publication date *(quality insurance/validation, overwritten regularly and to be used for user alerts, such as email or RSS feeds)* *DD-MM-YYYY*	**Mandatory**
Next publication date (3 fields) *-Year* ***YYYY*** *- Day (when available)* ***DD*** *- Month (when available)* ***MM*** ***Not relevant for static one-off dataset***	**Optional**
Periodicity: (e.g. *Annual, Monthly, Quarterly, etc.)*T*o be expressed in number of months associated to a label* (e.g. *1 month and label "Monthly", 12 months/label ""annual"; 18 months and label "18 months", etc.)* ***Not relevant for static one-off and dynamic dataset***	**Optional**
Update method - Static one-off - Static versioned - Dynamic *A dataset may be published with different update methods.* *Examples:* *- dynamic, updated continuously* *- static versioned, updated as a whole on a regular basis* *- static one-off, updated once only*	**Mandatory**

Dataset metadata	Status
Languages (ISO Codes) of the content - English - and/or French - and/or other language *Dataset are composed of both data and statistical metadata. Statistical Metadata exist in both English and French. The language of a dataset relates to the language in which titles, column and row headings, labels and notes are displayed.* *Currently, for most of the dataset contents there are:* *• One dataset in English, being the combination of data and statistical metadata in English* *• One dataset in French, being the combination of data and statistical metadata in French*	**Mandatory**
Size Number of cells (algorithm to calculate size to be defined)	**Optional**
Related countries *When the dataset is country specific, this relates to the Country ISO Code of the related country*	**Optional**
Period covered: Start year	**Optional**
Period covered: End year	**Optional**
Time range to be always formatted YYYY-YYYY for datasets (where the first year is the start year and the last year is the end year)	**Optional**
Variable index *(concept of statistics "variable" to present content on the index of statistical variables)* - A classification with variable titles in both English and French, although the French titles link to English content	**Optional**
Short abstracts *(in all available languages)*	**Optional**
Long abstracts *(in all available languages)*	**Optional**
Keyword(s) *(in all available languages) when available*	**Optional**
JEL classification *(in all available languages) when available* *- order (when several JEL classifications)* *- multiple JEL classifications are possible and they are ordered by relevance*	**Optional**
Theme(s) *(in all available languages) when available* *- classification into themes (theme classifications which may be internal or external to OECD website)*	**Optional**
Image file *Different types of images may be associated to a dataset (for displaying a brand image associated to a dataset)* *Images can be attached at the level of the rendition of a given format* e.g.***- type: print screen of web page*** *delivering dataset* ***- type: top banner for*** *stand-alone dataset* *Note: as for book covers different sizes of the same image may be required for specific channels.*	**Optional**
Year of publication	**Mandatory**

Links	Status
Belongs to (parent-child relationship): link to the statistical collection it belongs to ***Not relevant for stand-alone dataset*** **Attribute of link: order number** of the dataset within the statistical collection (managed by Editorial) --> used to display the dataset in the statistical collection table of contents.	**Mandatory** for non stand-alone dataset
Has main parent: link from dataset to its main parent collection (as opposed to other "step" parent collections) (relevant for datasets belonging to more than one collection)	**Optional**
Has physical form: link to content in various output formats/media: (MS Excel®, PDF, URL, external link…). Relates a rendition and its physical items *- PDF, MS Excel®, CSV format relate to a dataset filename.* *- online media may be a URL*	**Mandatory**
Is related to *Relates complementary products (*e.g. *analytical material) sharing the same or similar topics, all of them being recorded into OECD's publishing catalogue.* *Relates concepts and products together.* **+ *It should be possible to order related links of a dataset.***	**Optional**
Has external links *Relates complementary products (*e.g. *related databases, international statistical classification manuals) which are externally published. These related products are not catalogued in OECD's publishing catalogue.* ***+ It should be possible to order external links of a dataset.***	**Optional**
Has renditions Relates a product and its content at rendition level. *One rendition corresponds to a specific language version of a content. So this link allows one to relate the same content to several language versions.*	**Optional**
Supersedes *Indicates that a concept or product has been changed but the content remains the same. Can be used for table versions.*	**Optional**
Is continued by *Indicates that a product has been discontinued or is no longer available, and is replaced by a new one. Links products together.*	**Optional**
Is imprinted by *Indicates the legal/organisational body who owns the publication at imprint level.* - A logo in different sizes is associated to each imprint. Note: as for book covers different sizes of the same logo may be required for specific channels.	**Optional**
Is copyrighted by *Indicates legal / organisational body who owns the publication at copyright level.* By default OECD	**Mandatory**
Is edited by: for each potential owner a percentage of ownership will have to be assigned (*e.g.* PAC 25%, STD 30%, etc.). *Indicates the legal/organisational body who owns the publication at editorial responsibility level.* For each potential owner a percentage of ownership will have to be assigned.	**Mandatory**

Annex 2. Collection of datasets metadata

Definition	
3 types of collections 1) **Collection of datasets** which belongs to a parent collection **(=sub-collection)** * Does not have ISSN - Is NOT subject to subscription * Has a DOI (DOI suffix built as: <CollectionAcronym>-data) 2) **Collection of datasets** which does not belong to a parent collection * Has an ISSN * Has a DOI (DOI suffix built as: <CollectionAcronym>-data) *** Is subject to subscription** 3) **Collection of sub-collections** (and in some cases also datasets) * Has an ISSN * Has a DOI (DOI suffix = <CollectionAcronym>-data) * Is subject to subscription	Main Metadata exist at 3 levels: - conceptual - rendition - physical - This object is searchable on the iLibrary. - This object is not cited: *Statistical Collections are not cited as such, but the collection main title is included in the dataset citation - see dataset.*

Publishing metadata	**Status**
ISSN ***Not relevant for sub-collections of datasets (case 1 above)*** OECD statistical collections are by definition all available in electronic format (vs. print format). However, collections in electronic format do not necessarily contain data in one unique medium but may contain datasets in different media (multi-dimensional online dataset, PDF, MS Excel® , etc…). The ISSN is assigned for a collection in online medium in a given language and may contain datasets in different online media (multi-dimensional online dataset PDF, MS Excel®, etc…)	**Optional**
DOI *(digital object identifier of the content)* Each collection in one given language should have its own DOI. - The DOI of the English collection should resolve to the library homepage of the collection on the English interface - The DOI of the French collection should resolve to the library homepage of the collection on the French interface	**Mandatory**
Main title *(in all available languages)*	**Mandatory**
Subtitle *(in all available languages)*	**Optional**
Acronym *> required for DOI syntax (see in above definition of collection)*	**Mandatory**
Update method - Static stand-alone - Static versioned - Dynamic *A database collection may be published with different update methods.*	**Mandatory**

Publishing metadata	Status
Language(s) (ISO Codes) of the content - English - and/or French - and/or other language *Relates to the language in which titles, column and row headings, labels and notes are displayed.* *Similarly to datasets, 2 distinct collections in a given language can be defined for the same collection content depending on the languages of its datasets:* *• One collection in English, composed of datasets or sub-collections in English* *• One collection in French, composed of datasets or sub-collections in French*	**Mandatory**
Short abstracts *(in all available languages)*	**Optional**
Long abstracts: Label to display is "About (Hide/Show)" *(in all available languages)*	**Optional**
Keyword(s) *(in all available languages) when available*	**Optional**
Theme(s) *(in all available languages) when available* *- classification into themes on various websites*	**Optional**
JEL classification *(in all available languages) when available* *- multiple JEL classifications are possible and they are ordered by relevance*	**Optional**
Image file *Different types of images may be associated to statistical product:* *Images can be attached at the level of the rendition of a given format* e.g. *- type: print screen of web page delivering collection* *- type: top banner brand image for the statistical database collection*	**Optional**

Links	Status
Belongs to (parent-child relationship): link to the statistical collection it belongs to ***> relevant for sub-collection without ISSN belonging to a top collection*** **Attribute of link: Order number** of the sub-collection within the parent collection used to display the sub-collections of the top collection table of contents.	**Optional**
Has main parent : link from sub-collection to its main parent top collection (as opposed to other "step" parent collections) ***> relevant only for sub-collection belonging to more than one collection***	**Optional**
Has default child: link to the dataset or sub-collection for which it is the default child - A dataset/collection may be the "default" dataset/sub-collection of one or many collections	**Optional**
Has renditions *Relates a product and its content at rendition level.* *One rendition corresponds to a specific language version of a content. So this link allows one to relate the same content to several language versions.*	**Optional**
Is related to *Relates complementary products (*e.g. *analytical material), sharing the same or similar topics, all of them being recorded into OECD's publishing catalogue..* *Relates concepts together and products together.* ***+ It should be possible to order related links of a collection.***	**Optional**
Has external links *Relates the dataset collection to published items which are external to OECD (i.e. not in the OECD publishing catalogue).* ***+ It should be possible to order external links of a collection.***	**Optional**

Links	Status
Supersedes *Indicates that a concept or product has been changed but the content remains the same. Can be used for table versions.*	**Optional**
Is continued by *Indicates that a product has been discontinued or is no longer available, and is replaced by a new one. Links products together.*	**Optional**
Is imprinted by *Indicates the legal/organisational body who owns the publication at imprint level* - A logo in different sizes is associated to each imprint. Note: as for book covers, different sizes of the same logo may be required for specific channels	**Optional**
Is edited by For each potential owner a percentage of ownership will be assigned. . *Indicates the legal/organisational body who owns the publication at editorial responsibility level.*	**Derived from datasets for display**

Annex 3. Key table collection metadata

Definition of key table collection

Statistical collection related to a theme or to countries used for key tables publishing (*e.g.* Taxation: Key tables form OECD; OECD Country Statistical Profiles).

Main metadata exist at 1 level: conceptual. This object is not cited.

This object is searchable on the iLibrary

DOI suffix= ISSN

Publishing metadata	Status
ISSN *Each language version of a Key table collection edition has a distinct ISSN*	**Mandatory**
Collection type (code of key table country collection or key table thematic collection) *Used to distinguish type of collection for online publishing and citation rules*	**Mandatory**
DOI (digital object identifier of the content) *Each language version of a Key table collection has a distinct DOI.*	**Mandatory**
Collection Main Title *(in all available languages)*	**Mandatory**
Collection Subtitle *(in all available languages)*	**Optional**
Update method - Static stand-alone - Static versioned - Dynamic *A key table collection may be published with different update methods.*	**Mandatory**
Language(s) (ISO Codes) of the content: EN and maybe another language - English - and/or French - and/or other language *Relates to the language in which titles, column and row headings, labels and notes are displayed.*	**Mandatory**
Short abstracts *(in all available languages)*	**Mandatory**
Long abstracts *(in all available languages)*	**Mandatory**
Keyword(s) *(in all available languages) when available*	**Mandatory**
Theme(s) *(in all available languages) when available*	**Mandatory**

Links	
Has renditions *Links to the same content in another language,* *Relates a product and its content at rendition level*	**Mandatory**
Is related to *Relates complementary products (*e.g. *analytical material), sharing the same or similar topics, all of them being recorded into Kappa* *Relates concepts together and products together*	**Optional**
Has external links *Relates Kappa objects with external links*	**Optional**
Supersedes *Indicates that a concept or product has been changed but the content remains the same. Can be used for table versions.*	**Optional**
Is continued by *Indicates that a product has been discontinued or is no longer available, and is replaced by a new one. Links products together.*	**Optional**
Is edited by For each potential owner a percentage of ownership will have to be assigned. *Indicates the legal/organisational body who owns the publication at editorial responsibility level.*	**Optional**

Annex 4. Key tables metadata

Definition of key table

Defines table content singled out as very key interest at content and rendition level

This object is not cited.
This object is searchable on the iLibrary.

DOI suffix syntax:
<Key table collection ISSN>-table<Key table OrderNumber>
e.g. 16097319-table1

Note that Country Statistical Profile tables do not have a Key Table Order Number, but the ISO3 country acronym is used to identify each table:

DOI suffix syntax:
<Key table collection ISSN>-table-<ISO3>
e.g. 16097319-table-aus
(OECD Country Statistical Profile of Australia)

Publishing metadata	Status
DOI (digital object identifier of the content) *Each language version of a Key table has a distinct DOI.*	**Mandatory**
Main title *(in all available languages)*	**Mandatory**
Subtitle *(in all available languages)*	**Optional**
Periodicity May be daily, weekly, monthly, etc. To be expressed in number of days associated to a label	**Mandatory**
Update method - Static stand-alone - Static versioned - Dynamic *A key table series may be published with different update methods.*	**Mandatory**
Language(s) (ISO Codes) of the content: - English - and/or French - and/or other language *Relates to the language in which titles, column and row headings, labels and notes are displayed*	**Mandatory**
Source Note Free area of text when there is no “source” expressed as "Is Sourced From" metadata (*e.g.* when the data is calculated by author based on an unpublished database). *Source Note is language qualified (i.e. one note in English and one note in French can be supplied)* *This metadata is mutually exclusive with the metadata link "IsSourceFrom"*	**Optional**
Short abstracts *(in all available languages)*	**Mandatory**

Publishing metadata	Status
Long abstracts *(in all available languages)*	**Mandatory**
Keyword(s) *(in all available languages) when available*	**Optional**
JEL classification *(in all available languages) when available* *- multiple JEL classifications are possible and they are ordered by relevance*	**Optional**
Variable Index A classification with variable titles in both English and French, although the French titles link to English content.	**Optional**
Related Countries *When the key table is country specific, this relates to the Country ISO Code of the related country*	**Optional**
Theme(s): A thematic classification (*e.g.* 1 unique theme) *(in all available languages) when available*	**Optional**

Links	Status
Belongs to (parent-child relationship) Link to the key table collection it belongs to *Relates component concept to parent concept, or component product to parent product (at ISBN and ISSN level)* **Attribute of link** Order number of the key table within the table collection--> used to order the display of the tables in the key table collection table of contents. ***Not relevant for country profiles key tables***	**Mandatory**
Is related to *Relates complementary products (*e.g. *analytical material), sharing the same or similar topics, all of them being recorded into OECD publishing catalogue* *Relates concepts together and products together.*	**Optional**
Has external links *Relates complementary products (such as publications) sharing the same or similar topics, which are published external to OECD (i.e. not recorded in OECD publishing catalogue)*	**Optional**
Has renditions *Links to the same content in another language* *Relates a product and its content at rendition level*	**Optional**
Is sourced from A database (*e.g.* can be a dataset or a publication. If both exist as references, the database is the source while the publication is related to the table via the "Is Part Of" link. *Relates a product to the source(s) of the product* *This metadata link is mutually exclusive with the metadata "SourceNote"*	**Optional**
Supersedes *Indicates that a concept or product has been changed but the content remains the same. Can be used for table versions.*	**Optional**
Is continued by *Indicates that a product has been discontinued or is no longer available, and is replaced by a new one. Links products together.*	**Optional**
Is imprinted by *Indicates the legal/ organisational body who owns the publication at imprint level.* - A logo in different sizes is associated to each imprint.	**Optional**

Links	Status
Note: as for book covers different sizes of the same logo may be required for specific channels	
Is copyrighted by By default OECD *Indicates legal / organisational body who owns the publication at copyright level.*	**Mandatory**
Is edited by: For each owner a percentage of ownership will be assigned. *Indicates the legal/organisational body who owns the publication at editorial responsibility level*.	**Mandatory**

Annex 5. Key table edition metadata

Definition of key table edition

Defines an edited table edition (physical representations of a table as a two dimensional object) singled out as very key interest
Main metadata exist at 3 levels: conceptual, rendition, physical

This object is not searchable on the OECD iLibrary.

Citation rule:
<author physical/institutional> (<year of publication date>), "<Key tableTitle >", <*Key table collection title*>, No. <Key table Order number>.
doi: <doiprefix>/<Key table edition doisuffix>
<doi URL>
(Accessed on <date>)

OECD (2009), "Income tax plus employee social security contributions", *OECD Key Tables on Taxation,* No.1.
doi: 10.1787/16097319-2009-1- table1
http://dx.doi.org/10.1787/16097319-2009-1- table1
(Accessed on 21 January 2009)

Example Country Statistical Profiles citation:
OECD (2008), OECD Country Statistical Profile of Australia.
doi: 10.1787/16097319-2008-table-aus
http://dx.doi.org/10.1787/16097319-2008-table-aus
(Accessed on 21 December 2008)

DOI suffix syntax:
<Key table collection ISSN>-<EditionYear><Year of publication>-table<Key table OrderNumber>
e.g. 16097319-2009-table1

Note that Country Statistical Profile tables do not have a KeyTableOrderNumber, but the ISO3 country acronym is used to identify each table:

DOI suffix syntax for Country Statistical Profile:
<Key table collection ISSN>-<EditionYear>-table-<ISO3>
e.g. 16097319-2009-table-aus

Publishing metadata	Status
DOI (digital object identifier of the content) *Each language version of a Key table edition has a distinct DOI.*	**Mandatory**
Main title *(in all available languages)*	**Mandatory**
Subtitle *(in all available languages)*	**Optional**
Author(s) * Institutional author will be OECD unless external * Physical author - first name - last name - affiliation - order (when several authors - managed by OECD)	**Optional**

Publishing metadata	Status
Publication date (quality insurance/validation, overwritten regularly and to be used for email alerts) DD-MM-YYYY	**Mandatory**
Update method - Dynamic - Static The full-text items of a current *key table edition are continuously updated so "dynamic" and overwritten on line until a new key table edition is created. At this time, the full text item of the previous edition becomes "static"*	**Optional**
Language(s) (ISO Codes) of the content: EN and maybe another language - English - and/or French - and/or other language *Relates to the language in which titles, column and row headings, labels and notes are displayed*	**Mandatory**
Related countries *When the key table is country specific, this relates to the Country ISO Code of the related country.*	**Optional**
Edition year Each Keytable edition has a year edition	Mandatory
Year number To distinguish Key table editions released during the same year.	Mandatory
Period covered: Start year To be used for years structured as follows: yyyy : a single year yyyy-yyyy: a range of years (EVERY year between the two specified values) > it is mutual exclusive with time range	**Optional**
Period covered: End year To be used for years structured as follows: yyyy: a single year yyyy-yyyy: a range of years (EVERY year between the two specified values) > it is mutual exclusive with time range	**Optional**
Time range > It is mutual exclusive with period covered start and end year To be used for years structured as follows: […], yyyy: the previous range AND the specified values. Examples: **1980, 1995-2003** means: 1980, 1995, 1996, 1997, 1998, 1999, 2000, 2001, 2002, and 2003. **1970, 1980, 1990, 2000-2005** means: 1970, 1980, 1990, 2000, 2001, 2002, 2003, 2004, and 2005.	**Optional**

Links	Status
Is direct member of (parent-child relationship) Link to the key table it belongs to	**Mandatory**
Has physical form Link to content in various output formats/media: (MS Excel® , PDF, HTML external link…) Relates a rendition and its physical items - ***Is not subject to access rights***	**Mandatory**
Has renditions Relates a product and its content at rendition level One rendition corresponds to a specific language version of a content. So this link allows one to relate the same content to several language versions.	**Optional**
Supersedes Indicates that a concept or product has been changed but the content remains the same. Can be used for table versions.	**Optional**
Is edited by For each owner a percentage of ownership will have to be assigned. *Indicates the legal/organisational body who owns the publication at editorial responsibility level.*	**Optional**