

Índice de contenido

INTRODUCTION. BACKGROUND.....	5
Mission.....	5
Users	5
Where to find	5
DIMENSIONS OF ANALYSIS.....	6
Legal framework.....	7
Stage 1. Copyright.....	7
Stage 2. Private use.....	7
Stage 3. Non-commercial reuse.....	7
Stage 4. Commercial reuse.....	7
Stage 5. Only recognition.....	7
Technical Standards for data releasing.....	8
Stage 1. Closed standard	8
Stage 2. Open standard.....	8
Stage 3. Open standard with explanatory documentation.....	8
Stage 4. Open standards, individual access/metadata	8
Stage 5. Open standard with associated metadata and explanatory material.....	8
Accessibility to information.....	9
Stage 1. No web access or manual request.....	9
Stage 2. Web Access URL with registration or multiple sources.....	9
Stage 3. Web access or unique URL parameters.....	9
Stage 4. Web Access URL only or update parameters.....	9
Stage 5. API or query language.....	9
ASSESSING HOW TO.....	11
Example	11
MAPPING W3C 5 STARS AND MELODA.....	12
FAQ.....	13
THANKS TO.....	14
To do	15
Anex I. Meloda for Journalists (M4J).....	16



Change	Date	Description
Inclusion of M4J	23-10-11	Meloda for journalists (M4J) include as Anex I
Updata with economical terms to M4J	28-11-11	Meloda for journalists (M4J) include economical terms



INTRODUCTION. BACKGROUND

Mission

The aim of **MELODA** (**ME**thodology for re**L**easing **O**pen **DA**ta) is to provide a tool that will accelerate the release of information to society, mainly from the public sector but also from the private sphere.

Its use is focused on the maximum use of released information including commercial uses, mixing with private sources, etc. So that legal issues about information are considered as well as technical and accessibility issues.

Users

This methodology is aimed at those public and private entities that are releasing data and want to maximize the use of data published under a free reuse scheme.

Where to find

We release new versions of this methodology in <http://gobernamos.com> check it out.



DIMENSIONS OF ANALYSIS

To be operational **MELODA** must obtain a figure by assessing the available information of a data source (the same information that any user could reach). This requirement restricts the number of dimensions of analysis.

In this version **MELODA** analyses three dimensions:

- Legal framework
- Technical Standards
- Accessibility to the information

For each dimension to be analyzed 5 stages of maturity are considered.



Legal framework

In this five stages other considerations (i.e. Restriction of malicious use of information, misuse, requirements to include the last update's date, etc are not included).

Attribution to the original source is always considered as a must.

Stage 1. Copyright.

Data sources in this stage:

- Either will reserve the copyright of data, thus restricting unauthorized use
- Either those entities require of an non-online authorisation to use the information (i.e. A form that has to be manually answered fits this category).

Stage 2. Private use

Data sources in this stage will allow the use of the data without approval processes but only for private uses. Not public use available.

Stage 3. Non-commercial reuse

Data sources in this stage will allow reuse of data but they do not allow commercial uses of the information reuse.

(i.e. For a blog, for a non-profit organisation)

Stage 4. Commercial reuse

Data sources in this stage will allow reuse of data, including the commercial reuse.

(i.e. For any company in order to create new visualization of data)

Stage 5. Only recognition

Data sources in this stage will only ask re-users the attribution of the data.

i.e. Releasing data with CC BY license.



Technical Standards for data releasing

Stage 1. Closed standard

Data sources in this stage are release on proprietary standards. Other open formats but not suitable for reutilization will be considered closed as well. The definition of open standard can be found in national level legislation.

Eg xls, pdf, doc, shp, etc.

Stage 2. Open standard

Data sources in this stage are published on open standards but as individual files.

i.e. eg, csv, txt, odb, odt, ods, etc.

Stage 3. Open standard with explanatory documentation

Data sources in this stage includes those which release information as open standards in individual files but with available explanatory information about the contents of the files.

i.e. csv, odb, odt, ods, etc with additional information about size, data type, range of registers of the file.

Stage 4. Open standards, individual access/metadata

Data sources in this stage includes those which release information accessible item by item through some technical mechanism (API, individual URI, etc). Additional information about the content of the data source is not available.

i.e. rdf without additional instructions

Stage 5. Open standard with associated metadata and explanatory material

Data sources in this stage includes those which release information accessible on individual basis through some open technical mechanism (open API, individual URI, etc). Additional information about the content of the data source is available.

I.e. rdf with additional documentation



Accessibility to information

Stage 1. No web access or manual request

Access to information requires a non-automatic approval process for access to data information or to register data in a manual form.

Stage 2. Web Access URL with registration or multiple sources

Access to information via the web, but requires user interaction to select the data source.

Stage 3. Web access or unique URL parameters

Access to information via the web, but allows each of the datasets to be accessed individually, or through a shortened URL, or by specific parameters in the query call.

Stage 4. Web Access URL only or update parameters

Access to information via the web, but allows each of the datasets to be accessed individually, or through a shortened URL, or by specific parameters in the query call and includes the date of last data update.

Stage 5. API or query language

Access to information provides access to specific data of the dataset, either by calling a documented API or through a query language of data sources

(I.e. SPARQL).





ASSESSING HOW TO

Stages 1 to 5 for every dimension (legal, technical and accessibility) are marked as 0, 25, 50,75 and 100%

Every dimension is weighted equal than the others. So that the global mark for an individual datasource is composed by one third of every mark.

Global assessing of an institution that releases data is the average of the mark of every datasource released by this institution.

Example

Institution A releases an opendata source that you are assessing. It is released as a csv file without further instructions about its content only for private use through a unique URL.

Legal assessing : Stage 2 → 25%

Technical assessing: Stage 2 → 25%

Accessibility assessing: Stage 3 → 50%

global mark $(25\% + 25\% + 50\%) / 3 = 33,3\%$

The institution A which releases previous example has 2 additional sources apart from the above mentioned, one marked with 58% and another one marked 16%.

So that the final **MELODA** assessing for the institution is:

global mark $(33,3\% + 58\% + 16\%) / 3 = 35,8\%$



MAPPING W3C 5 STARS AND MELODA

5 stars	MELODA			
	MELODA	Legal	Technical	Accessibility
★ make your stuff available on the web (whatever format) but with an open licence	> 17%	Minimum 25 % It could be protected with an open licence but only private use	Minimum 0 % It could be privative standards	Minimum 25 % The access to information could require a manual selection of dataset although being in an URL
★★ Available as machine-readable structured data (e.g. excel instead of image scan of a table)	> 17%	Minimum 25 % It could be protected with an open licence but only private use	Minimum 0 % It could be privative standards	Minimum 25 % The access to information could require a manual selection of dataset although being in an URL
★★★ as (2) plus non-proprietary format (e.g. CSV instead of excel)	>25%	Minimum 25 % It could be protected with an open licence but only private use	Minimum 25 % It could be an non-proprietary format but without explanation about the content	Minimum 25 % The access to information could require a manual selection of dataset although being in an URL
★★★★ All the above plus, Use open standards from W3C (RDF and SPARQL) to identify things, so that people can point at your stuff	>67%	Minimum 25 % It could be protected with an open licence but only private use	100 % Using RDF and SPARQL	Minimum 75 % It could no tinclude specific information about how to use data
★★★★★ All the above, plus: Link your data to other people's data to provide context	>75%	Minimum 50 % It could be protected with an open licence but only for non commercial uses	100 % Using RDF and SPARQL	Minimum 75 % It could no tinclude specific information about how to use data



FAQ

Q: An entity have several datasets, How could we calculate the global assessment?

A: This is one of the goals of the methodology, just average all data source marks.

Q: The same dataset is released in several standards and licenses. Which should be the final mark?

A: The highest stage for every area (legal, technical or accessibility), bearing in mind that this situation could accessed by the user simultaneously.



THANKS TO

Igor Zubiarre the best collaborator in the world.

Francesco Cardi, for his great contributions to the debate.

Pro bono publico a Spanish organisation which hosts the debates to create MELODA.

Grupo de periodismo de datos del Medialab Prado

Mar Cabra

Access info Europe



To do

Check if it would possible to include into MELODA's next versions to assess the 'amount' of information an entity is releasing compared with its potential.



Anex I. Meloda for Journalists (M4J)

Meloda for journalists is a simplified version of Meloda methodology only focused in a primary use of the information from a qualitative point of view.

The simplified conditions are:

- 1) Data source has to be data. Not documents, not images unless they were attached to some other layers of data information.
- 2) Data has to be released in an open standard. If you need to pay for a tool to access the information it creates an economical barrier to the use.
- 3) Data has to be legally reusable. Data needs a legal framework which permits the commercial use of the data (CC by, PDDL, etc) are examples of approved licenses.
- 4) Data has to be accessible through a unique URL. Otherwise if manual interaction is needed reuse of the information needs of human capital and it constitutes an economical barrier.
- 5) Last but not least, access to the information has to be free (no cost).

The goal of this simplified version is to help reusers to know if a data source will be really helpful for a simplified full use

