EBI: <https://www.ebi.ac.uk/>

Controle de Variáveis:

|  |  |
| --- | --- |
| **URI used** | https://www.ebi.ac.uk/arrayexpress/experiments/E-GEOD-33601/ |
| **Analysis date** | 29/04/2020 |
| **Acronyms** | |
| **RaCE -** Researcher Compliance Experiment | |
| **RaCE -** Machine Compliance Experiment | |

FINDABLE

|  |  |  |
| --- | --- | --- |
| **Principle** | F1 | |
| **Description** | (Meta)data is assigned persistent and globally unique identifiers. | |
| **RaCE** | **/ \* Globally Unique? \* /**  The standardization, documentation, and use of URIs makes Identifiers globally unique.  **/ \* Persistent? \* /**  According to Re3Data the identifiers are persistent.  https://www.re3data.org/repository/r3d100010222  Identifiers are persistent when found on identifiers.org:  https://registry.identifiers.org/registry/arrayexpress  **/ \* Comments \* /**  Local identifiers are standardized and documented. | |
| **Result of RaCE (by color)** |  | |
| **Recommendations** | There are no improvements | |
| **MaCE** | **METRIC DESCRIPTION** | **ANALYZE** |
| **Unique Identifier**  Test whether the metadata resource has a unique identifier | It has an Uniform Resource Identifier type identifier . |
| **Identifier Persistence**  Metric to test whether the metadata resource's unique identifier is likely to be persistent. The known scheme is registered with FAIRSharing (https://fairsharing.org/standards/?q=&selected\_facets=type\_exact:identifier%20schema). For URLs that do not follow a scheme in FAIRSharing , we tested known URL persistence schemes ( purl , oclc , fdlp , purlz , w3id, ark ). | The persistence was found on Identifiers.org |
| **Data Identifier Persistence**  Metric to test whether the unique identifier for the data resource is likely to be persistent. The known scheme is registered with FAIRSharing (https://fairsharing.org/standards/?q=&selected\_facets=type\_exact:identifier%20schema). For URLs that do not follow a scheme in FAIRSharing , we tested known URL persistence schemes ( purl , oclc , fdlp , purlz , w3id, ark ). | The persistence was found on Identifiers.org |
| **Result of MaCE (by color)** | **Unique Identifier** |  |
| **Identifier Persistence** |  |
| **Data Identifier Persistence** |  |
| **Recommendations** | **Unique Identifier** | There are no improvements. |
| **Identifier Persistence** | There are no improvements. |
| **Data Identifier Persistence** | There are no improvements. |

|  |  |  |
| --- | --- | --- |
| **Principle** | F2 | |
| **Description** | Data is described with rich metadata. | |
| **RaCE** | **/ \* Is there metadata standardization? \* /**  Yes, the data are standardized.  **/ \* Is the metadata complete? \* /**  Perhaps, the metadata is based on the EFO ontology ( [https://www.ebi.ac.uk/ols/ontologies/efo](https://translate.google.com/translate?hl=pt-BR&prev=_t&sl=pt&tl=en&u=https://www.ebi.ac.uk/ols/ontologies/efo) ), not only for the ArrayExpress repository but all other EMBL-EBI repositories. Both intrinsic and contextual metadata are presented.  **/ \* Comments \* /**  There is not. | |
| **Result of RaCE (by color)** |  | |
| **Recommendations** | There are no improvements. | |
| **MaCE** | **Structured Metadata**  Tests whether a machine is capable of finding structured metadata. They can be (for example) RDFa , embedded json , json-ld or structured metadata negotiated by content, such as RDF Turtle . | RDFa- style metadata is found (analyzed via extruct. In general, EBI is well advanced in terms of providing metadata and structured data, there is a platform dedicated to linked data with RDF databases, and supporting SPARQL queries. |
| **Grounded Metadata**  Tests whether a machine is capable of finding grounded metadata. that is, metadata terms that are in a resolvable namespace, where the resolution leads to a definition of the meaning of the term. Examples include JSON-LD, embedded schema or any form of RDF. This test currently excludes XML, even when terms are spaced by name. Future versions of this test may be more flexible. | Yes, from the execution via extruct it was possible to achieve this metric. |
| **Result of MaCE (by color)** | **Structured Metadata** |  |
| **Grounded Metadata** |  |
| **Recommendations** | **Structured Metadata** | There are no improvements. |
| **Grounded Metadata** | There are no improvements. |

|  |  |  |
| --- | --- | --- |
| **Principle** | F3 | |
| **Description** | Metadata must clearly and explicitly include the data identifier described. | |
| **RaCE** | **/ \* Is it possible to identify the data in the metadata? How is it done? \* /**  By standardizing the repository's identifiers, the data is clearly identified in the metadata ( [https://www.ebi.ac.uk/arrayexpress/help/programmatic\_access.html](https://translate.google.com/translate?hl=pt-BR&prev=_t&sl=pt&tl=en&u=https://www.ebi.ac.uk/arrayexpress/help/programmatic_access.html) ) by using the / sample at the end of the URI it is possible to extract the given in question.  **/ \* Are there different identifiers for metadata and data? \* /**    There is not, it uses the same identifiers. However, to access the data, it is necessary to use a URI with increment of characters.    **/\*Comments\*/**    Reference other IDs external to the repository. | |
| **Result of RaCE (by color)** |  | |
| **Recommendations** | There are no improvements. | |
| **MaCE** | **Data Identifier Explicitly in Metadata**  Metric for testing whether metadata contains the data's unique identifier. This is done by looking for a variety of properties, including foaf : primaryTopic , schema : mainEntity , schema : distribution , sio: is-about and iao : is-about . The codeRepo sitory scheme is used for software versions . | It was not possible to find the data identifier in the metadata using any (common) property / predicate reserved for that purpose. |
| **Metadata Identifier Explicitly in Metadata**  Metric for testing whether metadata contains the unique identifier for the metadata itself. This is done using a variety of scraping tools , including resolving DOI metadata, using the Python tool ' extruct ' and other ... | From the extruct is extracted RDFa d page, it is also possible to collect the publication DOI. |
| **Result of MaCE (by color)** | **Data Identifier Explicitly in Metadata** |  |
| **Metadata Identifier Explicitly in Metadata** |  |
| **Recommendations** | **Data Identifier Explicitly in Metadata** | It is extremely important to use tags such as those mentioned in the metric, as it facilitates the use of (meta)data extraction tools.  To meet this metric, it is necessary to use knowledge representation languages. |
| **Metadata Identifier Explicitly in Metadata** | There are no improvements. |

|  |  |  |
| --- | --- | --- |
| **Principle** | F4 | |
| **Description** | (Meta)data is recorded or indexed in searchable resources. | |
| **RaCE** | **/ \* The URI or local ID returns the record in engines of search? \* /**  Yes, it is returned. | |
| **Result of RaCE (by color)** |  | |
| **Recommendations** | There are no improvements. | |
| **MaCE** | **Searchable in Major Search Engine**  It tests whether a machine is able to discover the resource by searching, using Google. | The searcher used the identifier and returned the record of (meta)data. |
| **Result of MaCE (by color)** | **Searchable in Major Search Engine** |  |
| **Recommendations** | **Searchable in Major Search Engine** | There are no improvements. |

ACCESSIBLE

|  |  |
| --- | --- |
| **Principle** | A1 |
| **Description** | (Meta)data is retrieved by its identifier using standardized communication protocols. |
| **RaCE** | **/ \* Do you use standardized protocols? \* /**    Yes, it uses standardized protocols for data access.    **/ \* Do you have proprietary software for data access? \* /**    It does not have any proprietary software.    **/\*Comments\*/**    It is possible to access via FTP. |
| **Result of RaCE (by color)** |  |
| **Recommendations** | There are no improvements. |
| **MaCE** | There is no FAIR METRICS GEN2 for this principle. |
| **Result of MaCE (by color)** | There is no FAIR METRICS GEN2 for this principle. |
| **Recommendations** | There is no FAIR METRICS GEN2 for this principle. |

|  |  |  |
| --- | --- | --- |
| **Principle** | A1.1 | |
| **Description** | Protocol is open, free and universally implemented. | |
| **RaCE** | **/ \* Does it have open, free and universally implemented access? \* /**    Yes, as long as the data is not private, it can be accessed without any need for identification or authorization.    **/ \* Comments \* /**    There isn't . | |
| **Result of RaCE (by color)** |  | |
| **Recommendations** | There are no improvements | |
| **MaCE** | **Uses Open Free Protocol for data Retrieval**  The data can be recovered by an open and free protocol. Tests the data GUID for your resolution protocol. Currently passes InChI keys , DOIs , identifiers and URLs . The recognition of other identifiers will be added at the request of the community. | Several identifiers are returned that link to the data, even those that are not in the repository. DOI, URL, proper identifiers are examples of return when using JSON. |
| **Uses Open Free Protocol for Metadata Retrieval**  The metadata can be retrieved by an open and free protocol. Tests the metadata GUID for your resolution protocol. Currently passes InChI keys , DOIs , identifiers and URLs . The recognition of other identifiers will be added at the request of the community. | Meets this metric has means of searches that prioritize XML with previously described tag and JSON fully identified. Returns DOI, URLs, Identifiers . |
| **Result of MaCE (by color)** | **Uses Open Free Protocol for data Retrieval** |  |
| **Uses Open Free Protocol for Metadata Retrieval** |  |
| **Recommendations** | **Uses Open Free Protocol for data Retrieval** | There are no improvements. |
| **Uses Open Free Protocol for Metadata Retrieval** | There are no improvements. |

|  |  |  |
| --- | --- | --- |
| **Principle** | A1.2 | |
| **Description** | The protocol allows for authentication and authorization when necessary. | |
| **RaCE** | **/ \* Does the protocol allow authentication and authorization? \* /**    Private data is generally pre- published / unpublished data. Access to private data is under password control.  You must follow the instructions to attend the means of authentication and authorization, present at: https://www.ebi.ac.uk/arrayexpress/help/programmatic\_access.html | |
| **Result of RaCE (by color)** |  | |
| **Recommendations** | There are no improvements | |
| **MaCE** | **Data Authentication and Authorization**  Test a discovered data GUID for the ability to implement authentication and authorization in your resolution protocol. Currently passes InChI keys , DOIs , identifiers and URLs . He also searches the metadata for the Dublin Core ' accessRights ' property , which can point to a document that describes the data access process. The recognition of other identifiers will be added at the request of the community. | The GUID returned from the metadata is a URI, known for allowing authentication / authorization. |
| **Metadata Authentication and Authorization**  Tests the metadata GUID for the ability to implement authentication and authorization in its resolution protocol. Currently passes InChI keys , DOIs , identifiers and URLs . The recognition of other identifiers will be added at the request of the community. | The GUID returned from the metadata is a URI, known to allow authentication / authorization. |
| **Result of MaCE (by color)** | **Data Authentication and Authorization** |  |
| **Metadata Authentication and Authorization** |  |
| **Recommendations** | **Data Authentication and Authorization** | There are no improvements. |
| **Metadata Authentication and Authorization** | There are no improvements. |

|  |  |  |
| --- | --- | --- |
| **Principle** | A2 | |
| **Description** | Metadata must be accessible even when data is no longer available. | |
| **RaCE** | **/ \* Is there a data versioning policy ? \* /**  It was not possible to identify.  **/ \* (Meta)data can be erased? \* /**  Yes, they can be deleted. It can be deleted under the following conditions: incorrect data, data published without permission of the legitimate owner, data published with change not made.  **/ \* (Meta)data can be updated? \* /**  Yes, they can be updated. Seen at: https: //www.ebi.ac.u k / arrayexpress / help / modify\_data\_loaded\_in\_AE.html  **/ \* Is it possible to access metadata of data that no longer exists? \* /**  Could not find documentation or examples.  **/ \* Is there a persistence policy ? \* /**  No persistence policy was found.  **/ \* Comments \* /**  There are no comments. | |
| **Result of RaCE (by color)** |  | |
| **Recommendations** | There are no improvements. | |
| **MaCE** | **Metadata Persistence**  Metric for testing whether metadata contains a persistence policy, explicitly identified by a key persistencePolicy (in hashed data) or by a http://www.w3.org/2000/10/swap/pim/doc#persistencePolicy predicate in linked data. | It was not possible to find any persistence policy using any approach |
| **Result of MaCE (by color)** | **Metadata Persistence** |  |
| **Recommendations** | **Metadata Persistence** | In order to comply with this metric, it is necessary to indicate a persistence policy in the registration source code. |

INTEROPERABLE

|  |  |  |
| --- | --- | --- |
| **Principle** | I1 | |
| **Description** | (Meta)data use formal, accessible, shared and widely applicable language for knowledge representation. | |
| **RaCE** | **/ \* Is there a use of languages ​​to represent knowledge? \* /**    It does not use any knowledge representation language.    **/ \* Taxonomies, ontologies, controlled vocabularies are referenced or found in the repository? (Must meet F 1) \* /**    Controlled vocabularies, thesaurus and general taxonomies are found, however not all meet the F1 principle.    **/ \* Comments \* /**    There isn't . | |
| **Result of RaCE (by color)** |  | |
| **Recommendations** | It is necessary to use some knowledge representation language (RDF, OWL, DAML-OIL, JSON LD, are highly used examples).  Taxonomies, ontologies, controlled vocabularies must comply with the F1 principle. | |
| **MaCE** | **Metadata Knowledge Representation Language (WEAK)**  Maturity indicator to test whether metadata uses a formal language widely applicable to knowledge representation. This particular test has a broad view of what defines a 'knowledge representation language'; in this assessment, anything that can be represented as structured data will be accepted. | This metric is reached through the execution of the extruct and through the extraction via JSON |
| **Metadata Knowledge Representation Language (STRONG)**  Maturity indicator to test whether metadata uses a formal language widely applicable to knowledge representation. This particular test has a broad view of what defines a 'knowledge representation language'; in this evaluation, a knowledge representation language is interpreted as one in which the terms are semantically based on ontologies. Any form of RDF will pass this test (including RDF that is automatically extracted by third party analyzers, such as Apache Tika ). | This metric is reached through the execution of the extruct and through the extraction via JSON |
| **Result of MaCE (by color)** | **Metadata Knowledge Representation Language (WEAK)** |  |
| **Metadata Knowledge Representation Language (STRONG)** |  |
| **Recommendations** | **Metadata Knowledge Representation Language (WEAK)** | There are no improvements. |
| **Metadata Knowledge Representation Language (STRONG)** | There are no improvements. |

|  |  |  |
| --- | --- | --- |
| **Principle** | I2 | |
| **Description** | (Meta)data use vocabularies that follow the FAIR principles. | |
| **RaCE** | **/ \* Do vocabularies follow FAIR principles? \* /**    They don't follow.    **/\*Comments\*/**    The controlled vocabulary is well documented and can be found in all instances of registration in the repository. | |
| **Result of RaCE (by color)** |  | |
| **Recommendations** | Vocabularies must meet all FAIR principles, as they can be “called” the base repository, vocabularies that meet FAIR principles facilitate the design of a repository based on FAIR principles.    Using FAIR DATA POINT can be a way to meet this principle. | |
| **MaCE** | **Metadata Uses Fair Vocabularies (WEAK)**  Maturity indicator to test whether linked data metadata uses resolved terms. This only test if they resolve, and FAIR data does not resolve, so it is a rather weak test. | It makes use of linked data. |
| **Uses Fair Vocabularies (STRONG) Metadata**  Maturity indicator to test whether linked data metadata uses resolved linked data (FAIR) terms. | No resolved predicates were found for linked data. |
| **Result of MaCE (by color)** | **Metadata Uses Fair Vocabularies (WEAK)** |  |
| **Metadata Uses Fair Vocabularies (STRONG)** |  |
| **Recommendations** | **Metadata Uses Fair Vocabularies (WEAK)** | There are no improvements. |
| **Metadata Uses Fair Vocabularies (STRONG)** | For linked data to be used, for this to occur, data must be written in knowledge representation languages.  Use of FAIR DATA POINTS would assist in meeting this metric. |

|  |  |  |
| --- | --- | --- |
| **Principle** | I3 | |
| **Description** | (Meta)data includes qualified references for other (meta)data. | |
| **RaCE** | **/ \* Is there use of structured data languages? \* /**    Uses RDFa, but they are not complete.    **/ \* Is there qualification among (meta)data entities? \* /**    It is found from the RDFa file, but they are not complete, they are not qualified.    **/ \* Comments \* /**    There isn't. | |
| **Result of RaCE (by color)** |  | |
| **Recommendations** | There are no improvements. | |
| **MaCE** | **Metadata Contains Qualified Outward References**  Maturity indicator to test whether metadata is externally linked to third-party resources. It only tests metadata that can be represented as linked data. | It is not provided with qualified references. |
| **Result of MaCE (by color)** | **Metadata Contains Qualified Outward References** |  |
| **Recommendations** | **Metadata Contains Qualified Outward References** | In order to be attended, it is necessary that the registry has a representation language that accepts the use of linked data. |

REUSABLE

|  |  |
| --- | --- |
| **Principle** | R1 |
| **Description** | (Meta)data are richly described with the plurality of precise and relevant attributes. |
| **RaCE** | **/ \* Is there a standardization of (meta)data? \* /**    Yes, the data is standardized for submission.    **/ \* Is there reference to metadata in other repositories? \* /**    Yes, there is an explicit reference to other external repositories    **/ \* Comments \* /**    There isn't. |
| **Result of RaCE (by color)** |  |
| **Recommendations** | There are no improvements. |
| **MaCE** | There is no FAIR METRICS GEN2 for this principle. |
| **Result of MaCE (by color)** | There is no FAIR METRICS GEN2 for this principle. |
| **Recommendations** | There is no FAIR METRICS GEN2 for this principle. |

|  |  |  |
| --- | --- | --- |
| **Principle** | R1.1 | |
| **Description** | (Meta)data is published with clear and accessible data usage licenses. | |
| **RaCE** | **/ \* Usage licenses are found in the registry? \* /**    Do not.    **/ \* Are usage licenses found in the repository? \* /**    Yes.    **/ \* Link to use licenses \* /**  [https://www.re3data.org/repository/r3d10 0011570](https://translate.google.com/translate?hl=pt-BR&prev=_t&sl=pt&tl=en&u=https://www.re3data.org/repository/r3d100011570)  [https://fairsharing.org/FAIRsharing.327nbg](https://translate.google.com/translate?hl=pt-BR&prev=_t&sl=pt&tl=en&u=https://fairsharing.org/FAIRsharing.327nbg) | |
| **Result of RaCE (by color)** |  | |
| **Recommendations** | The use licenses are found in the repository, however they are not referenced in the records. | |
| **MaCE** | **Metadata Includes License (Weak)**  Maturity indicator to test whether metadata contains an explicit license pointer. This test 'weak' will use a regular expression that is not case-sensitive and sweep the style metadata key / value, as well as metadad the linked data. Tests: xhtml , dvia , dcterms , cc , data.gov.au and Schema license predicates on linked data and validates the value of these properties. | Yes, this goal is met, there is a pointer to the license of the repository. terms - of -use is the identifier for the license |
| **Metadata Includes License (Strong)**  Maturity indicator to test whether the linked data metadata contains an explicit pointer to the license. Tests: license attributes xhtml , dvia , dcterms , cc , data.gov.au and Schema in linked data and validates the value of these properties. | Yes, this goal is met, there is a pointer to the license of the repository. terms - of -use is the identi |
| **Result of MaCE (by color)** | **Metadata Includes License (Weak)** |  |
| **Metadata Includes License (Strong)** |  |
| **Recommendations** | **Metadata Includes License (Weak)** | There are no improvements. |
| **Metadata Includes License (Strong)** | There are no improvements. |

|  |  |
| --- | --- |
| **Principle** | R1.2 |
| **Description** | (Meta)data are associated with detailed provenance. |
| **RaCE** | **/ \* Provenance (meta)data is identified? \* /**    Yes, the source metadata is identified.    **/ \* There are documents to describe the provenance \* /**    There are provenance documents for other EBI repositories ( [https://www.ebi.ac.uk/rdf/documentation/provenance/](https://translate.google.com/translate?hl=pt-BR&prev=_t&sl=pt&tl=en&u=https://www.ebi.ac.uk/rdf/documentation/provenance/) ), however for there isn’t for this. |
| **Result of RaCE (by color)** |  |
| **Recommendations** | Even if the (meta)data is associated with detailed provenance, there is no connection or documents that refer to the conception of a provenance. |
| **MaCE** | There is no FAIR METRICS GEN2 for this principle. |
| **Result of MaCE (by color)** | There is no FAIR METRICS GEN2 for this principle. |
| **Recommendations** | There is no FAIR METRICS GEN2 for this principle. |

|  |  |
| --- | --- |
| **Principle** | R1.3 |
| **Description** | (Meta)data meet community standards relevant to the domain. |
| **RaCE** | It is not possible to analyze this principle. |
| **Result of RaCE (by color)** | It is not possible to analyze this principle. |
| **Recommendations** | It is not possible to analyze this principle. |
| **MaCE** | There is no FAIR METRICS GEN2 for this principle. |
| **Result of MaCE (by color)** | There is no FAIR METRICS GEN2 for this principle. |
| **Recommendations** | There is no FAIR METRICS GEN2 for this principle. |