Document describing metrics for performing analyzes involving data repositories.

Description analysis

## Findable

|  |  |  |  |
| --- | --- | --- | --- |
| **Field** | **Description** | | |
| **Principle** | F1 | | |
| **Summary** | F1: (Meta) data are assigned globally unique and persistent identifiers | | |
| **How should it be analyzed?**  **(FAIR PRINCIPLES)** | 1st The (Meta) data must have some URI (Uniform Resource Identifier), such as:   * Compressed URI (CURIE) <prefix>: <local ID> * URIs (Prioritize what contains an identifier generated by the repository in the style <http uri pattern> <Local ID> - https://www.uniprot.org/uniprot/P98161) * URLs (In some cases, however, they must necessarily contain the Local ID)   2º The repository must contain some standardization in the design of its local identifiers, in order to guarantee that they are globally unique.  3º Analyze if the identifiers are persistent. For that, the persistence identifiers must be consulted:   * identifiers.org * https://github.com/perma-id/w3id.org * http://re3data.org/ (You must search the repository and thus analyze whether it has ID standardization and persistence) * Any other? | | |
| **How should it be analyzed?**  **(FAIR METRICS GEN2)** | **FM Gen2** | **Description** | **How is it being done?** |
| Unique Identifier | Testar se o recurso de metadados possui um identificador exclusivo. | Só testar se o URI é ou um URI compactado ou URI normal |
| Identifier Persistence | Métrica para testar se é provável que o identificador exclusivo do recurso de metadados seja persistente. O esquema conhecido é registrado no FAIRSharing (https://fairsharing.org/standards/?q=&selected\_facets=type\_exact:identifier%20schema). Para URLs que não seguem um esquema no FAIRSharing, testamos esquemas conhecidos de persistência de URL (purl, oclc, fdlp, purlz, w3id, ark). | Caso seja CURIE ver se o mesmo se encontra no FAIRSharing.  Caso seja URI, testar sua pesistência em:  PURL(<https://archive.org/services/purl/>),  OCLC (<http://bibpurl.oclc.org/maint/display.html>),  FDLP (Não consegui o link),  PURLZ (Mesmo que PURL),  W3ID (<https://github.com/perma-id/w3id.org>), ARK (<https://n2t.net/e/pub/naan_registry.txt> e <https://n2t.net/e/n2t_full_prefixes.yaml>) |
| Data Identifier Persistence | Métrica para testar se é provável que o identificador exclusivo do recurso de dados seja persistente. O esquema conhecido é registrado no FAIRSharing (https://fairsharing.org/standards/?q=&selected\_facets=type\_exact:identifier%20schema). Para URLs que não seguem um esquema no FAIRSharing, testamos esquemas conhecidos de persistência de URL (purl, oclc, fdlp, purlz, w3id, ark). | Caso seja CURIE ver se o mesmo se encontra no FAIRSharing.  Caso seja URI, testar sua pesistência em:  PURL(<https://archive.org/services/purl/>),  OCLC (<http://bibpurl.oclc.org/maint/display.html>),  FDLP (Não consegui o link),  PURLZ (Mesmo que PURL),  W3ID (<https://github.com/perma-id/w3id.org>), ARK (<https://n2t.net/e/pub/naan_registry.txt> e <https://n2t.net/e/n2t_full_prefixes.yaml>) |

|  |  |  |  |
| --- | --- | --- | --- |
| **Field** | **Description** | | |
| **Principle** | F2 | | |
| **Summary** | Dados são descritos com metadados ricos | | |
| **How should it be analyzed?**  **(FAIR PRINCIPLES)** | 1º Os metadados devem possuir descrições informativas de contexto, qualidade e condição, ou características de dados.  2º Deve ser possível encontrar o dado baseado nas informações contidas nos metadados, mesmo sem o identificador.  3º Dados intrínsecos (metadados capturados automaticamente), contextuais (metadados sobre o contexto que o dado foi coletado/gerado), ontologias, taxonomias, dicionários controlados, utilização de FAIR Data Point (FDP), fazem com que os metadados se tornem mais ricos. Caso siga FDP, levar em consideração a especificação em https://github.com/FAIRDataTeam/FAIRDataPoint-Spec/blob/master/spec.md | | |
| **How should it be analyzed?**  **(FAIR METRICS GEN2)** | **FM Gen2** | **Description** | **How is it being done?** |
| Structured Metadata | Testa se uma máquina é capaz de encontrar metadados estruturados. Podem ser (por exemplo) RDFa, json incorporado, json-ld ou metadados estruturados negociados por conteúdo, como RDF Turtle. | É utilizado o script em python baseado na biblioteca Extruct que retorna qualquer um destes formatos.  Deve ser retornado metadados estruturados do tipo key/value.  Deve ser retornado metadados estruturados do tipo Linked Data. |
| Grounded  Metadata | Testa se uma máquina é capaz de encontrar metadados do tipo 'grounded'. ou seja, termos de metadados que estão em um espaço para nome resolvível, em que a resolução leva a uma definição do significado do termo. Os exemplos incluem JSON-LD, embedded schema ou qualquer forma de RDF. Atualmente, este teste exclui o XML, mesmo quando os termos são espaçados por nome. Versões futuras deste teste podem ser mais flexíveis. | Deve ser encontrado algum tipo de Linked Data.  É utilizado o script em python baseado na biblioteca Extruct que retorna qualquer um destes formatos. |

|  |  |  |  |
| --- | --- | --- | --- |
| **Field** | **Description** | | |
| **Principle** | F3 | | |
| **Summary** | Metadados devem incluir claramente e explicitamente o identificador do dados descrito. | | |
| **How should it be analyzed?**  **(FAIR PRINCIPLES)** | 1º Analisar se o link para o dado é achável dentro dos metadados, para caso o dado esteja presente em outro URI.  2º Identificar alguma conexão entre metadados e dados padronizada. | | |
| **How should it be analyzed?**  **(FAIR METRICS GEN2)** | **FM Gen2** | **Description** | **How is it being done?** |
| Data Identifier Explicitly in Metadata | Métrica para testar se os metadados contêm o identificador exclusivo dos dados. Isso é feito procurando uma variedade de propriedades, incluindo foaf: primaryTopic, schema: mainEntity, schema: distribution, sio: is-about e iao: is-about. O esquema codeRepository é usado para versões de software. | Aplicando o script do Extruct e Web Page Inspector e analisar (manualmente) caso tenha algum tipo de propriedade pré definida. |
| Metadata Identifier Explicitly in Metadata | Métrica para testar se os metadados contêm o identificador exclusivo para os próprios metadados. Isso é feito usando uma variedade de ferramentas 'scraping', incluindo a resolução de metadados DOI, o uso da ferramenta Python 'extruct' e outras ... | Aplicando o script do Extruct e Web Page Inspector e analisar (manualmente) caso tenha algum tipo de propriedade pré definida. A partir do resultado, se positivo, analisar se possui ID exclusivos, CURIE, DOIs entre outros. |

|  |  |  |  |
| --- | --- | --- | --- |
| **Field** | **Description** | | |
| **Principle** | F4 | | |
| **Summary** | (Meta)dados são registrados ou indexados em recursos pesquisáveis | | |
| **How should it be analyzed?**  **(FAIR PRINCIPLES)** | 1º A partir do URI e/ou ID local pesquisar em algum recurso pesquisável, neste caso Google. | | |
| **How should it be analyzed?**  **(FAIR METRICS GEN2)** | **FM Gen2** | **Description** | **How is it being done?** |
| Searchable in Major Search Engine | Test whether a machine is able to discover the resource by searching, using Google | Feito manualmente. |

## Accessible

|  |  |  |  |
| --- | --- | --- | --- |
| **Field** | **Description** | | |
| **Principle** | A1 | | |
| **Summary** | (Meta)dados são recuperados por seu identificador usando protocolos de comunicação padronizados | | |
| **How should it be analyzed?**  **(FAIR PRINCIPLES)** | 1º Não deve se encontrar barreiras de softwares ou protocolos proprietários para acesso aos (meta)dados a partir de seus identificadores | | |
| **How should it be analyzed?**  **(FAIR METRICS GEN2)** | **FM Gen2** | **Description** | **How is it being done?** |
| Don't have FM Gen2 | Don't have FM Gen2 | Don't have FM Gen2 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Field** | **Description** | | |
| **Principle** | A1.1 | | |
| **Summary** | Protocolo é aberto, grátis e universalmente implementado. | | |
| **How should it be analyzed?**  **(FAIR PRINCIPLES)** | 1º Analisar se há possibilidade de o dado ser privado, isso em todo repositório.  2º Tentar fazer download de (meta)dados e ver se há alguma restrição.  3º Analisar se há meios de acesso aos (meta)dados via FTP ou outro meio para facilitação de Machine-Readable. | | |
| **How should it be analyzed?**  **(FAIR METRICS GEN2)** | **FM Gen2** | **Description** | **How is it being done?** |
| Uses Open Free Protocol for Data Retrieval | Os dados podem ser recuperados por um protocolo aberto e gratuito. Testa o GUID de dados para o seu protocolo de resolução. Atualmente passa keys InChI, DOIs, identificadores e URLs. O reconhecimento de outros identificadores será adicionado a pedido da comunidade. | Usando o script do Extruct juntamente com Web Page Inspector para analisar se algum URI, CURIE é retornado. |
| Uses Open Free Protocol for Metadata Retrieval | Os metadados podem ser recuperados por um protocolo aberto e gratuito. Testa o GUID de metadados para o seu protocolo de resolução. Atualmente passa keys InChI, DOIs, identificadores e URLs. O reconhecimento de outros identificadores será adicionado a pedido da comunidade. | Usando o script do Extruct juntamente com Web Page Inspector para analisar se algum URI, CURIE é retornado. |

|  |  |  |  |
| --- | --- | --- | --- |
| **Field** | **Description** | | |
| **Principle** | A1.2 | | |
| **Summary** | O protocolo permite por uma autenticação e autorização quando necessário. | | |
| **How should it be analyzed?**  **(FAIR PRINCIPLES)** | 1º Analisar se há possibilidade de o dado ser privado, isso em todo repositório. | | |
| **How should it be analyzed?**  **(FAIR METRICS GEN2)** | **FM Gen2** | **Description** | **How is it being done?** |
| Data Authentication and Authorization | Teste o GUID de dados descoberto para a capacidade de implementar autenticação e autorização em seu protocolo de resolução. Atualmente passa keys InChI, DOIs, identificadores e URLs. Ele também pesquisa os metadados pela propriedade 'accessRights' do Dublin Core, que pode apontar para um documento que descreve o processo de acesso a dados. O reconhecimento de outros identificadores será adicionado a pedido da comunidade. | É utilizado o script do extruct com web page inspector e se algum URI ou CURIE é retornado. Também é analisado se há alguma propriedade do tipo acessRights. |
| Metadata Authentication and Authorization | Testa o GUID de metadados quanto à capacidade de implementar autenticação e autorização em seu protocolo de resolução. Atualmente passa keys InChI, DOIs, identificadores e URLs. O reconhecimento de outros identificadores será adicionado a pedido da comunidade. | É utilizado o script do extruct com web page inspector e se algum URI ou CURIE é retornado. |

|  |  |  |  |
| --- | --- | --- | --- |
| **Field** | **Description** | | |
| **Principle** | A2 | | |
| **Summary** | Metadados devem ser acessíveis mesmo quando os dados não estão mais disponíveis. | | |
| **How should it be analyzed?**  **(FAIR PRINCIPLES)** | 1º Analisar se há a possibilidade dos (meta)dados ser alterado.  2º Caso haja ver se há algum tipo de versionamento dos (meta)dados.  3º Tentar acessar versões antigas (antigos IDs) para ver se é possível encontrar metadados.  4º Analisar se há metadados referenciando (meta)dados atualizados. | | |
| **How should it be analyzed?**  **(FAIR METRICS GEN2)** | **FM Gen2** | **Description** | **How is it being done?** |
| Metadata Persistence | Métrica para testar se os metadados contêm uma política de persistência, identificada explicitamente por uma key persistencePolicy (em dados com hash) ou por um http://www.w3.org/2000/10/swap/pim/doc#persistencePolicy predicado em linked data. | É utilizado o script do extruct com web page inspector e se algum dos identificadores requisitados. |

## Interoperable

|  |  |  |  |
| --- | --- | --- | --- |
| **Principle** | I1 | | |
| **Summary** | (Meta)dados usam linguagem formal, acessível, compartilhada e amplamente aplicável para representação do conhecimento. | | |
| **How should it be analyzed?**  **(FAIR PRINCIPLES)** | 1º Analisar se há vocabulários controlados comuns, ontologias, thesaurus(contendo identificadores persistentes e globalmente únicos).  2º Analisar se há alguma padronização de (meta)dados para garantir a interoperabilidade (RDF, JSON-LD, DAML-OIL, OWL) | | |
| **How should it be analyzed?**  **(FAIR METRICS GEN2)** | **FM Gen2** | **Description** | **How is it being done?** |
| Metadata Knowledge Representation Language (WEAK) | Indicador de maturidade para testar se os metadados usam uma linguagem formal amplamente aplicável à representação do conhecimento. Esse teste em particular tem uma visão ampla do que define uma 'linguagem de representação do conhecimento'; nesta avaliação, qualquer coisa que possa ser representada como dados estruturados será aceita. | É utilizado o script do extruct com web page inspector e se algum tipo de dado estruturado é encontrado atende este requisito. |
| Metadata Knowledge Representation Language (STRONG) | Indicador de maturidade para testar se os metadados usam uma linguagem formal amplamente aplicável à representação do conhecimento. Esse teste em particular tem uma visão ampla do que define uma 'linguagem de representação do conhecimento'; nesta avaliação, uma linguagem de representação do conhecimento é interpretada como aquela em que os termos são semanticamente fundamentados em ontologias. Qualquer forma de RDF será aprovada neste teste (incluindo RDF que é extraído automaticamente por analisadores de terceiros, como o Apache Tika). | É utilizado o script do extruct com web page inspector e se achar algum CURIE faz um check se o mesmo se encontra no FAIRSharing. |

|  |  |  |  |
| --- | --- | --- | --- |
| **Field** | **Description** | | |
| **Principle** | I2 | | |
| **Summary** | (Meta)dados usam vocabulários que seguem os princípios FAIR. | | |
| **How should it be analyzed?**  **(FAIR PRINCIPLES)** | 1º Analisar se os documentos para padronização dos (meta)dados obedecem os Princípios FAIR.  2º Analisar se há alguma similaridade ou utiliza FAIR Data Point. | | |
| **How should it be analyzed?**  **(FAIR METRICS GEN2)** | **FM Gen2** | **Description** | **How is it being done?** |
| Metadata Uses Fair Vocabularies (WEAK) | Indicador de maturidade para testar se os metadados de linked data usam termos resolvidos. Isso testa apenas se eles resolverem, e não se resolverem dados FAIR, portanto, é um teste um tanto Poor. | É utilizado o script do extruct com web page inspector para analisar se há ocorrência de linked data. |
| Metadata Uses Fair Vocabularies (STRONG) | Indicador de maturidade para testar se os metadados de linked data usam termos resolvidos para linked data (FAIR). | É utilizado o script do extruct com web page inspector e se algum tipo de dado estruturado encontrado atende este requisito.  Não foi realizado com Apache Tika |

|  |  |  |  |
| --- | --- | --- | --- |
| **Field** | **Description** | | |
| **Principle** | I3 | | |
| **Summary** | (Meta)data include qualified references to other (meta)data. | | |
| **How should it be analyzed?**  **(FAIR PRINCIPLES)** | 1º Analisar se consta algum tipo de qualificação entre as referências contidas nos (meta)dados, como consta em <http://www.uniprot.org/uniprot/C8V1L6.rdf> | | |
| **How should it be analyzed?**  **(FAIR METRICS GEN2)** | **FM Gen2** | **Description** | **How is it being done?** |
| Metadata Contains Qualified Outward References | Test whether metadata is externally linked to third-party resources. It only tests metadata that can be represented as linked data. | Extruct script and web page inspector are used. It is analyzed if there is a link with external resources. |

## Reusable

|  |  |  |  |
| --- | --- | --- | --- |
| **Field** | **Description** | | |
| **Principle** | R1 | | |
| **Summary** | (Meta)dados são descritos ricamente com a pluralidade de atributos precisos e relevantes. | | |
| **How should it be analyzed?**  **(FAIR PRINCIPLES)** | 1º Por ter muita similaridade com o F2 o mesmo também é subjetivo sobre o que significa em o (meta)dado ser ricamente descrito. Quanto mais descrições informativas conter melhor. | | |
| **How should it be analyzed?**  **(FAIR METRICS GEN2)** | **FM Gen2** | **Description** | **How is it being done?** |
| Don't have FM Gen2 | Don't have FM Gen2 | Don't have FM Gen2 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Field** | **Description** | | |
| **Principle** | R1.1 | | |
| **Summary** | (Meta)dados são publicados com licenças de uso de dados claras e acessíveis. | | |
| **How should it be analyzed?**  **(FAIR PRINCIPLES)** | 1º É analisado se há algum tipo de licença presente no registro do (meta)dado ou contida nas diretrizes do repositório.  2º É analisado, no Re3Data, se o repositório possui alguma licença de uso. | | |
| **How should it be analyzed?**  **(FAIR METRICS GEN2)** | **FM Gen2** | **Description** | **How is it being done?** |
| Metadata Includes License (WEAK) | Indicador de maturidade para testar se os metadados contêm um ponteiro explícito para a licença. Esse teste 'Poor' usará uma expressão regular que não diferencia maiúsculas de minúsculas e varrerá os metadados de estilo de key / valor, bem como os metadados de linked data. Testes: predicados de licença xhtml, dvia, dcterms, cc, data.gov.au e Schema nos linked data e valida o valor dessas propriedades. | É utilizado o script do extruct e web page inspector. É analisado se há ocorrência de algum direcionamento ou mensionamento a alguma licença de acesso. |
| Metadata Includes License (STRONG) | Indicador de maturidade para testar se os metadados de linked data contém um ponteiro explícito para a licença. Testes: predicados de licença xhtml, dvia, dcterms, cc, data.gov.au e Schema nos linked data e valida o valor dessas propriedades. | É utilizado o script do extruct e web page inspector. É analisado se há ocorrência para um ponteiro explícito de licença de uso. |

|  |  |  |  |
| --- | --- | --- | --- |
| **Field** | **Description** | | |
| **Principle** | R1.2 | | |
| **Summary** | (Meta)dados são associados a proveniência detalhada. | | |
| **How should it be analyzed?**  **(FAIR PRINCIPLES)** | 1º Analisar se há alguma citação sobre proveniência no registro do (meta)dado ou no repositório como um todo.  2º Analisar se há incidência de proveniências prospectiva e/ou retrospectiva. | | |
| **How should it be analyzed?**  **(FAIR METRICS GEN2)** | **FM Gen2** | **Description** | **How is it being done?** |
| Don't have FM Gen2 | Don't have FM Gen2 | Don't have FM Gen2 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Field** | **Description** | | |
| **Principle** | R1.3 | | |
| **Summary** | (Meta)dados atendem aos padrões das comunidades relevantes para o domínio | | |
| **How should it be analyzed?**  **(FAIR PRINCIPLES)** | 1º Necessário analisar se os (meta)dados são considerados relevantes pela comunidade. Como realizar isto?? | | |
| **How should it be analyzed?**  **(FAIR METRICS GEN2)** | **FM Gen2** | **Description** | **How is it being done?** |
| Don't have FM Gen2 | Don't have FM Gen2 | Don't have FM Gen2 |

Melhorias (Princípios FAIR e Métricas)

## Findable

### Human

#### F1

|  |  |  |  |
| --- | --- | --- | --- |
| Levels | Color Representation | Score | Criteria |
| Exemplary |  | 3 | Todos registros possuem identificadores padronizados (necessário haver documentação), são globalmente únicos (CURIE ou URI) e são persistentes (devem estar contidos em algum repositório de persistência ou garantir a sua persistência). |
| Average |  | 2 | Todos registros possuem identificadores padronizados (necessário haver documentação), são globalmente únicos (CURIE ou URI). Não são persistentes. |
| Poor |  | 1 | Todos os registros são globalmente únicos (CURIE ou URI) OU são padronizados. Não são persistentes. |

#### F2

|  |  |  |  |
| --- | --- | --- | --- |
| Levels | Color Representation | Score | Criteria |
| Exemplary |  | 3 | Metadados possuem descrições informativas de contexto, qualidade e condição, ou características de dados. Deve ser possível encontrar o dado a partir do metadado (mesmo sem identificador). Deve conter dados intrínsecos, contextuais (ontologias, taxonomias, dicionários controlados). Os metadados são padronizados. |
| Average |  | 2 | Metadados não são completos, falta algum tipo de metadado (contexto, qualidade, ou característica de dados). É possível encontrar o dado a partir do metadado. Os metadados não são totalmente padronizados. |
| Poor |  | 1 | Metadados não são completos e/ou não são padronizados. Não são encontrados metadados essenciais. |

#### F3

|  |  |  |  |
| --- | --- | --- | --- |
| Levels | Color Representation | Score | Criteria |
| Exemplary |  | 3 | Há um link para o dado dentro dos metadados, para caso o dado esteja presente em outro URI. É estabelecida uma identificação sobre a conexão entre metadados e dados sendo a mesma padronizada. Utilização de linguagens como RDF atendem a este critério, sendo baseadas em ontologias. |
| Average |  | 2 | O dado e/ou link do dado é encontrado dentro dos metadados, não há definição de uma padronização entre a conexão entre dados e metadados. |
| Poor |  | 1 | Não é possível localizar o dado e/ou link do dado dentro dos metadados, o mesmo pode ser até encontrado mas não é subjetivo. |

#### F4

|  |  |  |  |
| --- | --- | --- | --- |
| Levels | Color Representation | Score | Criteria |
| Exemplary |  | 3 | Os (meta)dados são pesquisáveis, a partir de seu ID exclusivo, em um indexador de sites conhecido. O indexador deve possuir informações adicionais e não somente link para os (meta)dados (o repositório deve permitir a execução de crawlers para atingir este objetivo). |
| Average |  | 2 | Os (meta)dados são pesquisáveis, a partir de seu ID exclusivo, em um indexador de sites conhecidos. Não há informações adicionais no indexador (não foi permitido a utilização de crawlers). |
| Poor |  | 1 | Os (meta)dados não são pesquisáveis, a partir de seu ID exclusivo, em um indexador de sites conhecido. |

### Machine

#### FM-Gen2-F1.1

|  |  |  |  |
| --- | --- | --- | --- |
| Levels | Color Representation | Score | Criteria |
| Exemplary |  | 3 | Os metadados possui um identificador exclusivo. |
| Average |  | 2 |  |
| Poor |  | 1 | Os metadados não possuem identificador exclusivo. |

#### FM-Gen2-F1.2

|  |  |  |  |
| --- | --- | --- | --- |
| Levels | Color Representation | Score | Criteria |
| Exemplary |  | 3 | O identificador exclusivo do recurso de metadados é persistente. |
| Average |  | 2 |  |
| Poor |  | 1 | O identificador exclusivo do recurso de metadados não é persistente. |

#### FM-Gen2-F1.3

|  |  |  |  |
| --- | --- | --- | --- |
| Levels | Color Representation | Score | Criteria |
| Exemplary |  | 3 | O identificador exclusivo do recurso de dados não é persistente. |
| Average |  | 2 |  |
| Poor |  | 1 | O identificador exclusivo do recurso de dados não é persistente. |

#### FM-Gen2-F2.1

|  |  |  |  |
| --- | --- | --- | --- |
| Levels | Color Representation | Score | Criteria |
| Exemplary |  | 3 | Há metadados estruturados. Podem ser (por exemplo) RDFa, json incorporado, json-ld ou metadados estruturados negociados por conteúdo, como RDF Turtle. |
| Average |  | 2 |  |
| Poor |  | 1 | Não há metadados estruturados. |

#### FM-Gen2-F2.2

|  |  |  |  |
| --- | --- | --- | --- |
| Levels | Color Representation | Score | Criteria |
| Exemplary |  | 3 | São encontrados metadados do tipo 'grounded'. Os exemplos incluem JSON-LD, embedded schema ou qualquer forma de RDF. |
| Average |  | 2 |  |
| Poor |  | 1 | Não são encontrados metadados do tipo ‘grounded’. |

#### FM-Gen2-F3.1

|  |  |  |  |
| --- | --- | --- | --- |
| Levels | Color Representation | Score | Criteria |
| Exemplary |  | 3 | Os metadados contém o identificador exclusivo dos dados. foaf: primaryTopic, schema: mainEntity, schema: distribution, sio: is-about e iao: is-about, são exemplos a serem encontrados. O esquema codeRepository é usado para versões de software. |
| Average |  | 2 |  |
| Poor |  | 1 | Não há identificador exclusivo dos dados. |

#### FM-Gen2-F3.2

|  |  |  |  |
| --- | --- | --- | --- |
| Levels | Color Representation | Score | Criteria |
| Exemplary |  | 3 | Os metadados contêm o identificador exclusivo para os próprios metadados. Resultado de ferramentas 'scraping', incluindo a resolução de metadados DOI, o uso da ferramenta Python 'extruct' e outras ... |
| Average |  | 2 |  |
| Poor |  | 1 | Os metadados não contêm identificador exclusivo para os próprios metadados. |

#### FM-Gen2-F4

|  |  |  |  |
| --- | --- | --- | --- |
| Levels | Color Representation | Score | Criteria |
| Exemplary |  | 3 | É capaz de descobrir o recurso por pesquisa, usando indexador de sites. |
| Average |  | 2 |  |
| Poor |  | 1 | Não é capaz de descobrir o recurso por pesquisa, usando indexador de sites. |

## Accessible

### Human

#### A1

|  |  |  |  |
| --- | --- | --- | --- |
| Levels | Color Representation | Score | Criteria |
| Exemplary |  | 3 | Não há barreiras de softwares ou protocolos proprietários para acesso aos (meta)dados a partir de seus identificadores. |
| Average |  | 2 | Há barreiras de softwares ou protocolos proprietários para acesso aos (meta)dados a partir de seus identificadores, mas os mesmos podem ser utilizados por todos usuários. |
| Poor |  | 1 | Há barreiras de softwares ou protocolos proprietários para acesso aos (meta)dados a partir de seus identificadores, contudo os mesmos não podem ser utilizados por todos usuários. |

#### A1.1

|  |  |  |  |
| --- | --- | --- | --- |
| Levels | Color Representation | Score | Criteria |
| Exemplary |  | 3 | Os (meta)dados são abertos e é possível realizar download dos mesmos sem restrições de autorização ou autenticação de usuário. Utilização de outras formas de acesso como FTP são diferenciais. |
| Average |  | 2 | Os (meta)dados são abertos e é possível realizar download dos mesmos, contudo são encontradas restrições para o download, como softwares proprietários. |
| Poor |  | 1 | Os (meta)dados são abertos e é possível realizar download dos mesmos, contudo não é possível realizar o download do dado seja por inúmeras razões, como por exemplo, uso de softwares proprietários. |

#### A1.2

|  |  |  |  |
| --- | --- | --- | --- |
| Levels | Color Representation | Score | Criteria |
| Exemplary |  | 3 | É necessário que haja etapas de autenticação e autorização para acesso aos dados. A descrição para atendimento a estas fases é bem documentada com critérios para autenticação e autorização. |
| Average |  | 2 | É necessário que haja etapas de autenticação e autorização para acesso aos dados. Não há documentação para o atendimento destas fases. |
| Poor |  | 1 | É necessário que haja etapas de autenticação e autorização para acesso aos dados, contudo a mesma apresenta falhas. |

#### A2

|  |  |  |  |
| --- | --- | --- | --- |
| Levels | Color Representation | Score | Criteria |
| Exemplary |  | 3 | O (meta)dados podem ser alterados. Há algum tipo de versionamento dos (meta)dados. Os IDs devem ser persistentes. É padronizado e documentado, como a utilização de políticas de persistência. |
| Average |  | 2 | O (meta)dados podem ser alterados. Há algum tipo de versionamento dos (meta)dados. Os IDs devem não são persistentes. Não é padronizado e documentado. |
| Poor |  | 1 | O (meta)dados podem ser alterados. Não há versionamento dos (meta)dados. Os IDs devem não são persistentes. Não é padronizado e documentado. |

### Machine

#### FM-Gen2-A1.1.1

|  |  |  |  |
| --- | --- | --- | --- |
| Levels | Color Representation | Score | Criteria |
| Exemplary |  | 3 | Os dados podem ser recuperados por um protocolo aberto e gratuito. Resultados como keys InChI, DOIs, identificadores e URLs são aceitos. |
| Average |  | 2 |  |
| Poor |  | 1 | Os dados não podem ser recuperados por um protocolo aberto e gratuito. |

#### FM-Gen2-A1.1.2

|  |  |  |  |
| --- | --- | --- | --- |
| Levels | Color Representation | Score | Criteria |
| Exemplary |  | 3 | Os metadados podem ser recuperados por um protocolo aberto e gratuito. Resultados como keys InChI, DOIs, identificadores e URLs são aceitos. |
| Average |  | 2 |  |
| Poor |  | 1 | Os metadados não podem ser recuperados por um protocolo aberto e gratuito. |

#### FM-Gen2-A1.2.1

|  |  |  |  |
| --- | --- | --- | --- |
| Levels | Color Representation | Score | Criteria |
| Exemplary |  | 3 | O GUID de dados descoberto é capaz de implementar autenticação e autorização em seu protocolo de resolução. Resultados como keys InChI, DOIs, identificadores e URLs são aceitos. Ele também pesquisa os metadados pela propriedade 'accessRights' do Dublin Core, que pode apontar para um documento que descreve o processo de acesso a dados. O reconhecimento de outros identificadores será adicionado a pedido da comunidade. |
| Average |  | 2 |  |
| Poor |  | 1 | O GUID de dados descoberto não é capaz de implementar autenticação e autorização em seu protocolo de resolução. |

#### FM-Gen2-A1.2.2

|  |  |  |  |
| --- | --- | --- | --- |
| Levels | Color Representation | Score | Criteria |
| Exemplary |  | 3 | O GUID de metadados descoberto é capaz de implementar autenticação e autorização em seu protocolo de resolução. Resultados como keys InChI, DOIs, identificadores e URLs são aceitos. Ele também pesquisa os metadados pela propriedade 'accessRights' do Dublin Core, que pode apontar para um documento que descreve o processo de acesso a dados. O reconhecimento de outros identificadores será adicionado a pedido da comunidade. |
| Average |  | 2 |  |
| Poor |  | 1 | O GUID de metadados descoberto não é capaz de implementar autenticação e autorização em seu protocolo de resolução. |

#### FM-Gen2-A2

|  |  |  |  |
| --- | --- | --- | --- |
| Levels | Color Representation | Score | Criteria |
| Exemplary |  | 3 | Os metadados contêm uma política de persistência, identificada explicitamente por uma key persistencePolicy (em dados com hash) ou por um http://www.w3.org/2000/10/swap/pim/doc#persistencePolicy predicado em linked data. |
| Average |  | 2 |  |
| Poor |  | 1 | Não há política de persistência explícita por uma key. |

## Interoperable

### Human

#### I1

|  |  |  |  |
| --- | --- | --- | --- |
| Levels | Color Representation | Score | Criteria |
| Exemplary |  | 3 | Há vocabulários controlados comuns, ontologias, thesaurus(contendo identificadores persistentes e globalmente únicos). Há padronização de (meta)dados que garante a interoperabilidade (RDF, JSON-LD, DAML-OIL, OWL) |
| Average |  | 2 | Há vocabulários controlados comuns, ontologias e/ou thesaurus(contendo identificadores persistentes e globalmente únicos). Não há padronização de (meta)dados que garante a interoperabilidade (RDF, JSON-LD, DAML-OIL, OWL) |
| Poor |  | 1 | Não há vocabulários controlados comuns, ontologias e/ou thesaurus(contendo identificadores persistentes e globalmente únicos). Não há padronização de (meta)dados que garante a interoperabilidade (RDF, JSON-LD, DAML-OIL, OWL) |

#### I2

|  |  |  |  |
| --- | --- | --- | --- |
| Levels | Color Representation | Score | Criteria |
| Exemplary |  | 3 | O vocabulário controlado é usado para descrever datasets são bem documentados e solucionáveis usando identificadores persistentes e globalmente únicos e este documento são facilmente usável e acessível por todos que usam o dataset. |
| Average |  | 2 | O vocabulário controlado é usado para descrever datasets não são tão bem documentados e solucionáveis usando identificadores persistentes e globalmente únicos e/ou este documento não é tão facilmente usável e acessível por todos que usam o dataset ou ainda pode não ser existente. |
| Poor |  | 1 | Não há definição de vocabulário controlado para descrever datasets usando identificadores persistentes e globalmente únicos e/ou este documento não é usável e acessível por todos que usam o dataset ou não existe. |

#### I3

|  |  |  |  |
| --- | --- | --- | --- |
| Levels | Color Representation | Score | Criteria |
| Exemplary |  | 3 | (meta)dados possuem qualificações entre eles, muito bem estabelecidas fundamentadas e representadas por meio de linguagens de representação do conhecimento (como RDF) e ontologias. |
| Average |  | 2 | Não há qualificação entre os (meta)dados, contudo há utilização de linguagens de representação do conhecimento (como RDF) |
| Poor |  | 1 | Não há qualificação entre as relações dos (meta)dados e não há utilização de linguagens de representação do conhecimento (como RDF). |

### Machine

#### FM-Gen2-I1.1

|  |  |  |  |
| --- | --- | --- | --- |
| Levels | Color Representation | Score | Criteria |
| Exemplary |  | 3 | Os metadados usam uma linguagem formal amplamente aplicável à representação do conhecimento. Qualquer coisa que possa ser representada como dados estruturados será aceita. |
| Average |  | 2 |  |
| Poor |  | 1 | Os metadados não usam uma linguagem formal amplamente aplicável à representação do conhecimento. |

#### FM-Gen2-I1.2

|  |  |  |  |
| --- | --- | --- | --- |
| Levels | Color Representation | Score | Criteria |
| Exemplary |  | 3 | Os metadados usam uma linguagem formal amplamente aplicável à representação do conhecimento. Nesta avaliação, uma linguagem de representação do conhecimento é interpretada como aquela em que os termos são semanticamente fundamentados em ontologias. Qualquer forma de RDF será aprovada neste teste (incluindo RDF que é extraído automaticamente por analisadores de terceiros, como o Apache Tika). |
| Average |  | 2 |  |
| Poor |  | 1 | Os metadados não usam uma linguagem formal amplamente aplicável à representação do conhecimento. |

#### FM-Gen2-I2.1

|  |  |  |  |
| --- | --- | --- | --- |
| Levels | Color Representation | Score | Criteria |
| Exemplary |  | 3 | Os metadados de linked data usam termos resolvidos. Isso testa apenas se eles resolverem, e não se resolverem dados FAIR, portanto, é um teste um tanto Poor. |
| Average |  | 2 |  |
| Poor |  | 1 | Os metadados de linked data não usam termos resolvidos. |

#### FM-Gen2-I2.2

|  |  |  |  |
| --- | --- | --- | --- |
| Levels | Color Representation | Score | Criteria |
| Exemplary |  | 3 | Os metadados de linked data usam termos resolvidos para linked data (FAIR). |
| Average |  | 2 |  |
| Poor |  | 1 | Os metadados de linked data não usam termos resolvidos para linked data (FAIR). |

#### FM-Gen2-I3

|  |  |  |  |
| --- | --- | --- | --- |
| Levels | Color Representation | Score | Criteria |
| Exemplary |  | 3 | Os metadados se vinculam externamente a recursos de terceiros. Apenas testa metadados que podem ser representados como linked data. |
| Average |  | 2 |  |
| Poor |  | 1 | Os metadados não se vinculam externamente a recursos de terceiros |

#### 

## Reusable

### Human

#### R1

|  |  |  |  |
| --- | --- | --- | --- |
| Levels | Color Representation | Score | Criteria |
| Exemplary |  | 3 | (Meta)dados são muito bem descritos com pluralidade de atributos. Deve haver padronização e documentação dos dados. |
| Average |  | 2 | (Meta)dados não são tão bem descritos com pluralidade de atributos. Pode haver alguma padronização e documentação dos dados. |
| Poor |  | 1 | (Meta)dados não descritos com pluralidade de atributos. Não há padronização e documentação dos dados. |

#### R1.1

|  |  |  |  |
| --- | --- | --- | --- |
| Levels | Color Representation | Score | Criteria |
| Exemplary |  | 3 | Faz uso de licenças de uso muito bem definidas e documentadas e que sejam possíveis do usuário achar as mesmas tanto no registro quanto no repositório. |
| Average |  | 2 | Faz uso de licenças de uso, mas não são tão claras e definidas. Pode ser que não sejam de fácil acesso ao usuário. |
| Poor |  | 1 | Não há especificação de licenças de uso tanto para o repositório quanto para o registro. |

#### R1.2

|  |  |  |  |
| --- | --- | --- | --- |
| Levels | Color Representation | Score | Criteria |
| Exemplary |  | 3 | Há especificação da proveniência prospectiva e retrospectiva, ela é bem detalhada e fundamentada, está presente nos (meta)dados. |
| Average |  | 2 | Não há especificação de proveniência por meio de documentos, mas é possível encontrar nos (meta)dados proveniência subjetivamente. |
| Poor |  | 1 | Não há documentação sobre proveniência, existem (meta)dados ligados a proveniência mas os mesmos não são corretos e/ou não são completos. |

#### R1.3

|  |  |  |  |
| --- | --- | --- | --- |
| Levels | Color Representation | Score | Criteria |
| Exemplary |  | 3 | Há confirmação e certificação da comunidade sobre os (meta)dados representados no repositório. |
| Average |  | 2 | Os (meta)dados não são certificados pela comunidade, mas representam (meta)dados consolidados por comunidades do meio. |
| Poor |  | 1 | Os (meta)dados não são certificados e nem representam (meta)dados consolidados por comunidades do meio. |

### Machine

#### FM-Gen2-R1.1.1

|  |  |  |  |
| --- | --- | --- | --- |
| Levels | Color Representation | Score | Criteria |
| Exemplary |  | 3 | Os metadados contêm um ponteiro explícito para a licença. Esse teste 'Poor' usará uma expressão regular que não diferencia maiúsculas de minúsculas e varrerá os metadados de estilo de key / valor, bem como os metadados de linked data. Testes: predicados de licença xhtml, dvia, dcterms, cc, data.gov.au e Schema nos linked data e valida o valor dessas propriedades. |
| Average |  | 2 |  |
| Poor |  | 1 | Os metadados não contêm um ponteiro explícito para a licença. |

#### FM-Gen2-R1.1.2

#### 

|  |  |  |  |
| --- | --- | --- | --- |
| Levels | Color Representation | Score | Criteria |
| Exemplary |  | 3 | Os metadados de linked data contém um ponteiro explícito para a licença. Testes: predicados de licença xhtml, dvia, dcterms, cc, data.gov.au e Schema nos linked data e valida o valor dessas propriedades. |
| Average |  | 2 |  |
| Poor |  | 1 | Os metadados de linked data não contém um ponteiro explícito para a licença. |

Referências:

Identifiers for the 21st century: How to design, provision, and reuse persistent identifiers to maximize utility and impact of life science data <https://www.biorxiv.org/content/biorxiv/early/2017/03/20/117812.full.pdf>