Key Recommendations to improve FAIR Principles adaptation (References listed provide specific details and more extensive discussion on these topics)

## Findable

#### F1

|  |  |  |
| --- | --- | --- |
| **Field** | **Recommendation** | **References** |
| **FAIR Principles** | A utilização de identificadores globalmente únicos e persistentes pode ser feita por meio da criação de modelos próprios ou pela utilização de modelos e serviços altamente consolidados. | <https://www.go-fair.org/fair-principles/f1-meta-data-assigned-globally-unique-persistent-identifiers/>  <https://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.2001414>  <https://www.sciencedirect.com/science/article/pii/S2666389920300040>  <https://www.sciencedirect.com/science/article/pii/S1567422311000251> |
| **FMGEN2 - Unique Identifier** | Utilização de formatos de identificadores no estilo URI (compactado ou normal) | <https://www.w3.org/Addressing/URL/uri-spec.html>  <https://www.ietf.org/rfc/rfc3986.txt>  <https://whatis.techtarget.com/definition/URI-Uniform-Resource-Identifier>  <https://link.springer.com/referenceworkentry/10.1007%2F978-1-4419-9863-7_1572>  <https://www.w3.org/TR/2010/NOTE-curie-20101216/>  <https://www.crossref.org/blog/curies-a-cure-for-uris/>  <https://www.crossref.org/blog/curies-a-cure-for-uris/> |
| **FMGEN2 - Identifier Persistence** | Foco na adoção de identificadores persistentes para os metadados armazenados, sendo esses identificadores criados pelo próprio repositório (obedecendo regras de persistência) ou utilizando esquemas conhecidos de persistência de identificadores. | Adicionar todos os links para registro de persistência  <https://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.2001414>  <https://www.sciencedirect.com/science/article/pii/S1567422311000251>  <https://www.sciencedirect.com/science/article/abs/pii/S1574954116300140>  <https://www.sciencedirect.com/science/article/pii/S1368837519303021>  <https://www.sciencedirect.com/science/article/abs/pii/S0167739X18314638> |
| **FMGEN2 - Data Identifier Persistence** | Foco na adoção de identificadores persistentes para os dados armazenados, sendo esses identificadores criados pelo próprio repositório (obedecendo regras de persistência) ou utilizando esquemas conhecidos de persistência de identificadores. | Adicionar todos os links para registro de persistência  <https://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.2001414>  <https://www.sciencedirect.com/science/article/pii/S1567422311000251>  <https://www.sciencedirect.com/science/article/abs/pii/S1574954116300140>  <https://www.sciencedirect.com/science/article/pii/S1368837519303021>  <https://www.sciencedirect.com/science/article/abs/pii/S0167739X18314638> |

#### F2

|  |  |  |
| --- | --- | --- |
| **Field** | **Recommendation** | **References** |
| **FAIR Principles** | Os metadados encontrados pelos usuários devem ser os mais completos possíveis. Controle na inserção de metadados deve ser alta principalmente para que não haja metadados perdidos quando o detentor do dado inserir ao repositório, uma política de upload de metadados é fundamental.  Criação de uma padronização de metadados fundamentada é considerada essencial nesta parte, utilização de ontologias, vocabulários, thesaurus já existentes ou a criação de novos auxilia no aumento da riqueza de metadados.  Informações sobre dados, como sua descrição, descrição de colunas, significado de abreviações, unidades de medida, figuras, análises sobre os dados em si, como os dados foram coletados e/ou sugestões em como reutilizar os dados são exemplo de metadados essenciais. Outro fator de recomendação é criação de documentações auxiliar os usuários detentores de dados conhecerem sobre os requisitos para submissão de (meta)dados e além disso, deixar o mais intuitivo possível a partir da utilização de processos já fundamentados na área de IHC (Interação Humano Computador). | <https://www.go-fair.org/fair-principles/f2-data-described-rich-metadata/>  <https://theodi.org/knowledge-opinion/guides/>  <https://isa-tools.org/>  <https://www.dtls.nl/fair-data/find-fair-data-tools/> |
| **FMGEN2 - Structured Metadata** | Usar dados estruturados que sejam compatíveis para execução de aplicações voltadas para o scraping de dados, linguagens como RDFa, json incorporado, json-ld ou metadados estruturados negociados por conteúdo, como RDF Turtle, são exemplos de linguagens que suportam scraping de dados.  Existem inúmeros meios de padronização de estrutura de metadados, onde os mesmos ficam localizados no código fonte da página, ou por meio de uma extensão e designada para os dados estruturados. |  |
| **FMGEN2 - Grounded Metadata** | Grounded metadata refere-se a utilização de padronização de dados por meio de linguagens reconhecidas por computadores para scrapping de dados e etc. Contudo neste caso a fim de aumentar ainda mais a automação destas aplicações é necessário a utilização de named nodes ao invés de blank nodes, ou seja, utilizar identificadores (persistentes) ao invés de caracteres (dbpedia é um bom exemplo onde encontrar identificadores persistentes para nomes *incluir mais*) | <https://books.google.com.br/books?id=oY1lAgAAQBAJ&pg=PT43&lpg=PT43&dq=grounded+rdf&source=bl&ots=6TMIKsE9_0&sig=ACfU3U1auvNc_TNB1jh78KvbxH9QYkyCWQ&hl=pt-BR&sa=X&ved=2ahUKEwilyLiAr5nqAhVIDrkGHTGzBY0Q6AEwAHoECAYQAQ#v=onepage&q=grounded%20rdf&f=false> |

#### F3

|  |  |  |
| --- | --- | --- |
| **Field** | **Recommendation** | **References** |
| **FAIR Principles** | É necessário explicitar os identificadores dos datasets (dados) aos metadados para que assim os mesmos possam ser conectados, mesmo de uma forma simples. Utilização de identificadores persistentes e globalmente únicos (F1) auxilia a atingir esse princípio.  Para aprimorar ainda mais seria essencial adotar meio para que se possa realizar extração de dados via automação. |  |
| **FMGEN2 - Data Identifier Explicitly in Metadata** | Necessário possuir algum estruturação de dados que seja reconhecida por máquina (e.g., RDFa, json incorporado, json-ld ou metadados estruturados negociados por conteúdo, como RDF Turtle), além disso, possuir especificações/propriedades/tags com o intuito de localizar o identificador do(s) dado(s) facilmente nos metadados. |  |
| **FMGEN2 - Metadata Identifier Explicitly in Metadata** | O quanto mais utilizar metadados resolvíveis (que utilizam URIs) melhor será para aplicações automática de coleta de dados entenderem sobre os (meta)dados presentes naquele repositório. Novamente a utilização de padronização de dados estruturados que seja reconhecida por máquinas melhor será o reuso. |  |

#### F4

|  |  |  |
| --- | --- | --- |
| **Field** | **Recommendation** | **References** |
| **FAIR Principles** | Uma das possíveis melhorias é a padronização dos identificadores locais que auxiliam em engines de busca, tendo em vista que a partir de seus crawlers eles identificariam os ID Locais bem como os URIs. |  |
| **FMGEN2 - Searchable in Major Search Engine** | Liberar acesso para crawlers, padronização e documentação de ID é essencial para atendimento desta métrica. |  |

## Accessible

#### A1

|  |  |  |
| --- | --- | --- |
| **Field** | **Recommendation** | **References** |
| **FAIR Principles** | Criar meios para que os (meta)dados possam ser acessados sem que haja necessidade de softwares e protocolos proprietários. Utilizar protocolos amplamente conhecidos (e.g. http, https, ftp ...) e não utilização de softwares e padrões de arquivos não proprietários (e.g. utilizar csv ao invés de arquivos em excel). Além disso, é visto a necessidade de documentar e demonstrar ao usuário como é processo para realizar o download de dados, bem como deixar o processo o mais intuitivo possível e acessível (IHC - Interação Humano Computador). |  |

#### A1.1

|  |  |  |
| --- | --- | --- |
| **Field** | **Recommendation** | **References** |
| **FAIR Principles** | No caso para acesso a dados que não possuam restrição de acesso é essencial que os mesmos sejam acessados por meio de protocolos não proprietários, não ofereçam nenhum tipo de restrição para acesso aos dados, forneça meios para entendimento de como os dados podem ser acessados sem restrição(e.g. documentos, tutoriais). |  |
| **FMGEN2 - Uses Open Free Protocol for Data Retrieval** | Utilização de URIs (e.g. keys InChI, DOIs, identificadores e URLs) dentro dos metadados e que retornam o dado propriamente dito sem necessidade de autenticação e autorização e uso de protocolos ou softwares proprietários. |  |
| **FMGEN2 - Uses Open Free Protocol for Metadata Retrieval** | Usar URIs (e.g. keys InChI, DOIs, identificadores e URLs) nos metadados que retornem informações a respeito dos metadados sem a necessidade de protocolos restritivos ou uso de softwares proprietários. |  |

#### A1.2

|  |  |  |
| --- | --- | --- |
| **Field** | **Recommendation** | **References** |
| **FAIR Principles** | Caso haja a necessidade de restrição dos (meta)dados (e.g. dados sensíveis ligados a saúde) é necessário estabelecer meios para que haja restrição dos (meta)dados por meio de autenticação do usuário por meio do repositório e autorização por meio do detentor do dado ou repositório. É de extrema importância que todo o processo para acesso ao dado seja documentado para o entendimento do usuário o qual necessita utilizar o dado em questão. |  |
| **FMGEN2 - Data Authentication and Authorization** | Fornecer meios para que seja possível realizar autenticação e autorização de forma automática para o acesso aos dados ou para que seja identificado que o usuário devem passar por etapas de autenticação e/ou autorização para acesso aos dados (e.g. utilizar tags em linguagens reconhecíveis por máquinas que identificam que há essa necessidade). |  |
| **FMGEN2 - Metadata Authentication and Authorization** | Deve-se levar em consideração a necessidade de fornecer algum tipo de metadado de livre acesso, com intuito de auxiliar o usuário em ter conhecimento no dado que o mesmo terá acesso, com a finalidade do usuário ter certeza de que é o dado que necessita assim podendo otimizar as funções tanto do lado do usuário quanto do lado do detentor do dado e do repositório. As mesmas recomendações aplicadas ao FMGEN2 - Data Authentication and Authorization se aplicam a essa recomendação. |  |

#### A2

|  |  |  |
| --- | --- | --- |
| **Field** | **Recommendation** | **References** |
| **FAIR Principles** | Criação de políticas de versionamentos (citar algumas existentes), utilização de identificadores globalmente únicos e persistentes com o armazenamento de versões antigas do metadados (fracas - as mais básicas, fortes - mais completas).  Metadados no tipo texto, link para o dado mais atualizado, metadados sobre o detentor do dado, link para repositório de versionamento contendo metadados antigos, são exemplo de melhoria.  Resolução de possíveis URIs anteriores aos (meta)dados para que o usuário não seja levado a erros no acesso (resolução para o retorno ao URI mais novo).  Ter atenção a este princípio principalmente por conta do F4 onde o URI antigo pode estar indexado com maior valor que o URI novo. |  |
| **FMGEN2 - Metadata Persistence** | Utilização de políticas de persistência e versionamento são essenciais para não gerar problemas ao (meta)dado (e.g. inconsistência), ao usar políticas neste nível é necessário que as mesmas sejam vistas por máquinas, para assim auxiliar no entendimento de onde encontrar o (meta)dado mais novo caso sejam acessadas versões antigas, ou que o mesmo não está mais disponível. |  |

## Interoperable

#### I1

|  |  |  |
| --- | --- | --- |
| **Field** | **Recommendation** | **References** |
| **FAIR Principles** | Prover meios para que os (meta)dados sejam detalhados ao máximo no que tange a utilização de linguagens de representação do conhecimento (e.g. RDF, OWL, DAML-OIL, JSON LD, são exemplos altamente utilizados) e utilização de tesauros, ontologias, vocabulários controlados, onde os mesmos atendam ao requisito F1. O atendimento a essas recomendações auxilia no aumento da interoperabilidade entre diferentes (meta)dados |  |
| **FMGEN2 - Metadata Knowledge Representation Language** | Utilização de linguagens de representação do conhecimento e fundamentar (meta)dados baseado em ontologias, thesauros, dicionários controlados onde os mesmos atendem aos princípios FAIR. |  |

#### I2

|  |  |  |
| --- | --- | --- |
| **Field** | **Recommendation** | **References** |
| **FAIR Principles** | Essa recomendação é atendida a partir da adaptação ou desenvolvimento de vocabulários utilizados pelos repositórios aos princípios FAIR, a padronização pode se dar por meio da adoção do FAIRification Process e adoção do FAIR Data Point. |  |
| **FMGEN2 - Metadata Uses Fair Vocabularies** | Essa recomendação é atendida a partir da adaptação ou desenvolvimento de vocabulários utilizados pelos repositórios aos princípios FAIR, a padronização pode se dar por meio da adoção do FAIRification Process e adoção do FAIR Data Point. |  |

#### I3

|  |  |  |
| --- | --- | --- |
| **Field** | **Recommendation** | **References** |
| **FAIR Principles** | Referências qualificadas são atingidas a partir da utilização de linguagens de representação do conhecimento e utilização de identificadores globalmente únicos e persistentes, sendo assim baseando em padronizações para obter grounded (meta)data e consequentemente qualificações entre as possíveis relações entre os (meta)data. |  |
| **FMGEN2 - Metadata Contains Qualified Outward References** | Utilização de URIs externos em linguagens de representação do conhecimento com tags de qualificação. |  |

## Reusable

#### R1

|  |  |  |
| --- | --- | --- |
| **Field** | **Recommendation** | **References** |
| **FAIR Principles** | Completude dos (meta)dados facilitam o seu reuso. Quanto mais diversificação de (meta)dados melhor para seu reuso. (Meta)dados envolvendo o captura, análise, refinamento, instruções de uso, processamento, softwares e hardwares utilizados são essenciais para o entendimento dos dados armazenados. Essencial haver discussões entre as partes interessadas sobre quais (meta)dados devem ser essenciais para que o dado seja publicado/armazenado. |  |

#### R1.1

|  |  |  |
| --- | --- | --- |
| **Field** | **Recommendation** | **References** |
| **FAIR Principles** | Usar e abusar em licenças de uso bem fundamentadas (e.g. Creative Commons Licenses) para determinar as responsabilidades dos reutilizadores de (meta)dados, principalmente para dados sensíveis/restritos. Caso os (meta)dados sejam de livre acesso os mesmos devem estar baseados em licenças de uso livre. |  |
| **FMGEN2 - Metadata Includes License** | Necessário incluir nos registros meios para que seja possível identificar as licenças de uso. Utilização de licença xhtml, dvia, dcterms, cc, data.gov.au e Schema nos linked data são exemplos para atender esta métrica. |  |

#### R1.2

|  |  |  |
| --- | --- | --- |
| **Field** | **Recommendation** | **References** |
| **FAIR Principles** | Utilização de proveniência tanto prospectiva quanto retrospectiva com foco na documentação do ciclo de vida de objetos digitais, utilização de modelos de proveniência já fundamentados para extrair quais dados são essenciais para o reutilizador, há sistemas de proveniência que facilitam ao atendimento desta recomendação. |  |

#### R1.3

|  |  |  |
| --- | --- | --- |
| **Field** | **Recommendation** | **References** |
| **FAIR Principles** | Aceitação por parte de padronização de comunidades relevantes (citar algumas comunidades gerais) ao tema do (meta)dado acaba por agregar diversos benefícios ao próprio (meta)dado e ao repositório. Necessário pesquisar quais padronizações existentes ao tema, e, possivelmente, implementar ao controle de dados a serem inseridos pelos detentores dos (meta)dados. |  |

#### 