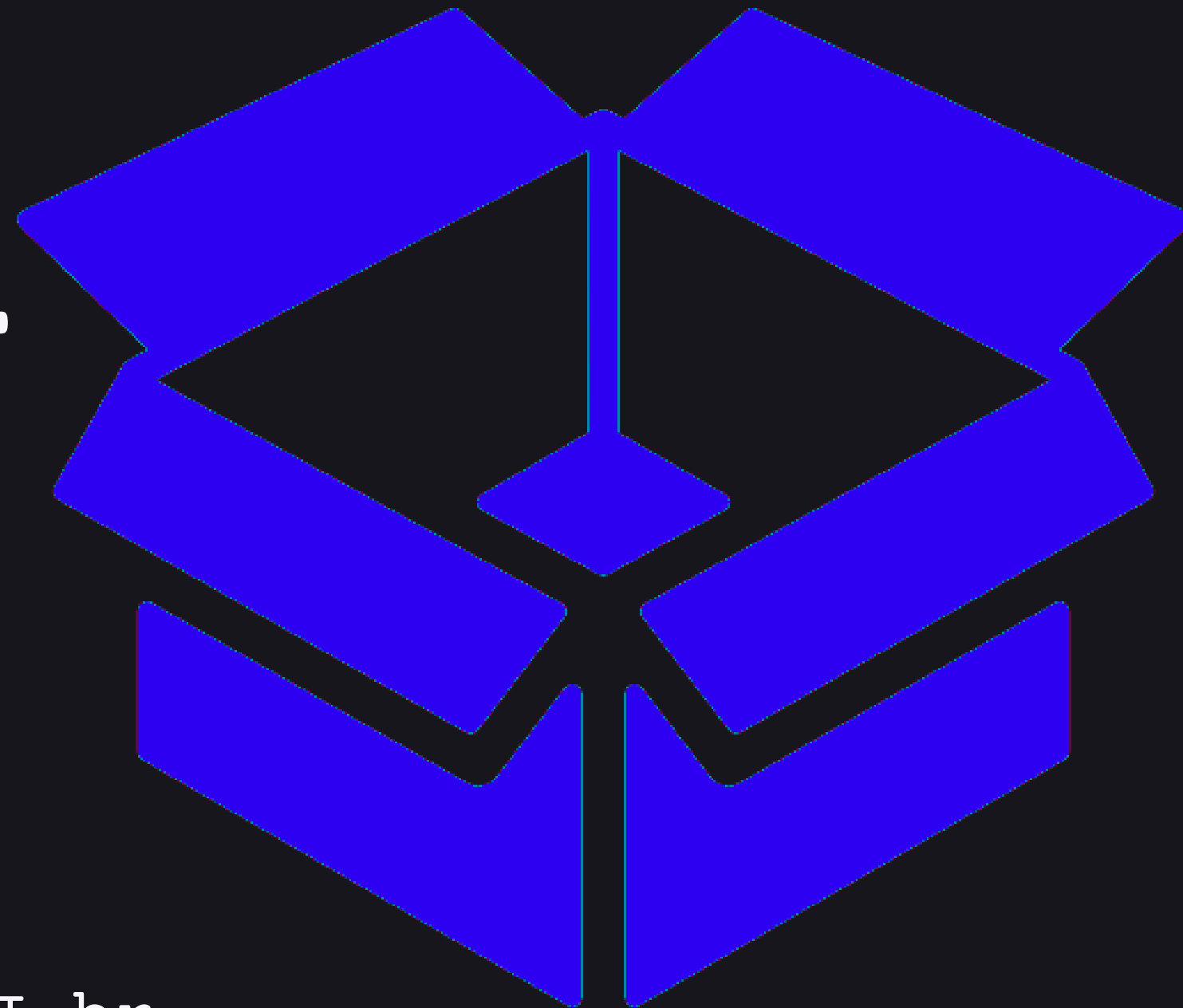


Dados Abertos e Boas Práticas para Dados na Web



ABRALIN EM CENA 17

Ana Eliza Duarte - Ceweb.br / NIC.br / CGI.br

São Paulo, 26 de junho de 2023

Alguns Conceitos



Dados



Informação



Dados
Abertos

1	0	1	1	0	0	1	0	1
1	1	0	1	1	0	0	1	0
1	0	1	0	0	1	1	0	1
1	1	1	1	0	1	1	0	1
1	0	1	0	0	1	0	1	0
1	1	0	1	1	1	1	0	1
1	0	1	1	0	1	0	1	0
1	1	0	1	1	1	1	0	0



"Dado é uma representação
quantificada e
quantificável."

(SETZER, 2004)

Dados



Informação é um conjunto
de dados compreendidos
por um receptor.

(SETZER, 2004)

Informação



Conjuntos de dados que podem ser utilizados, modificados e compartilhados por qualquer pessoa.

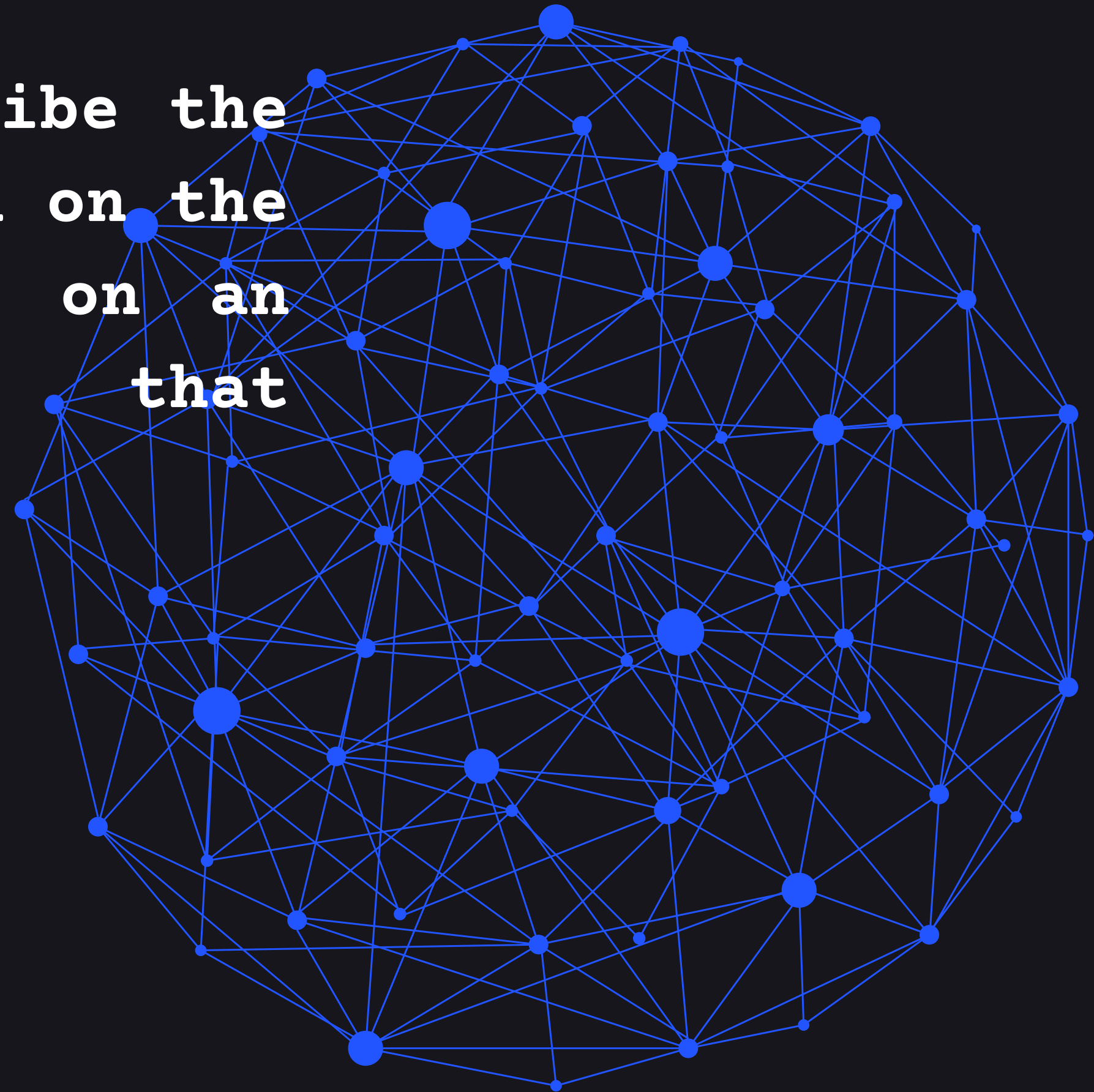
(Open Knowledge Foundation)

Dados Abertos

Web Semântica

The Semantic Web aims to describe the meaning of information published on the Web to enable retrieval based on an accurate understanding of that information's semantics"

RODRÍGUEZ-ROCHA et al., 2015)



5 Star Linked Data

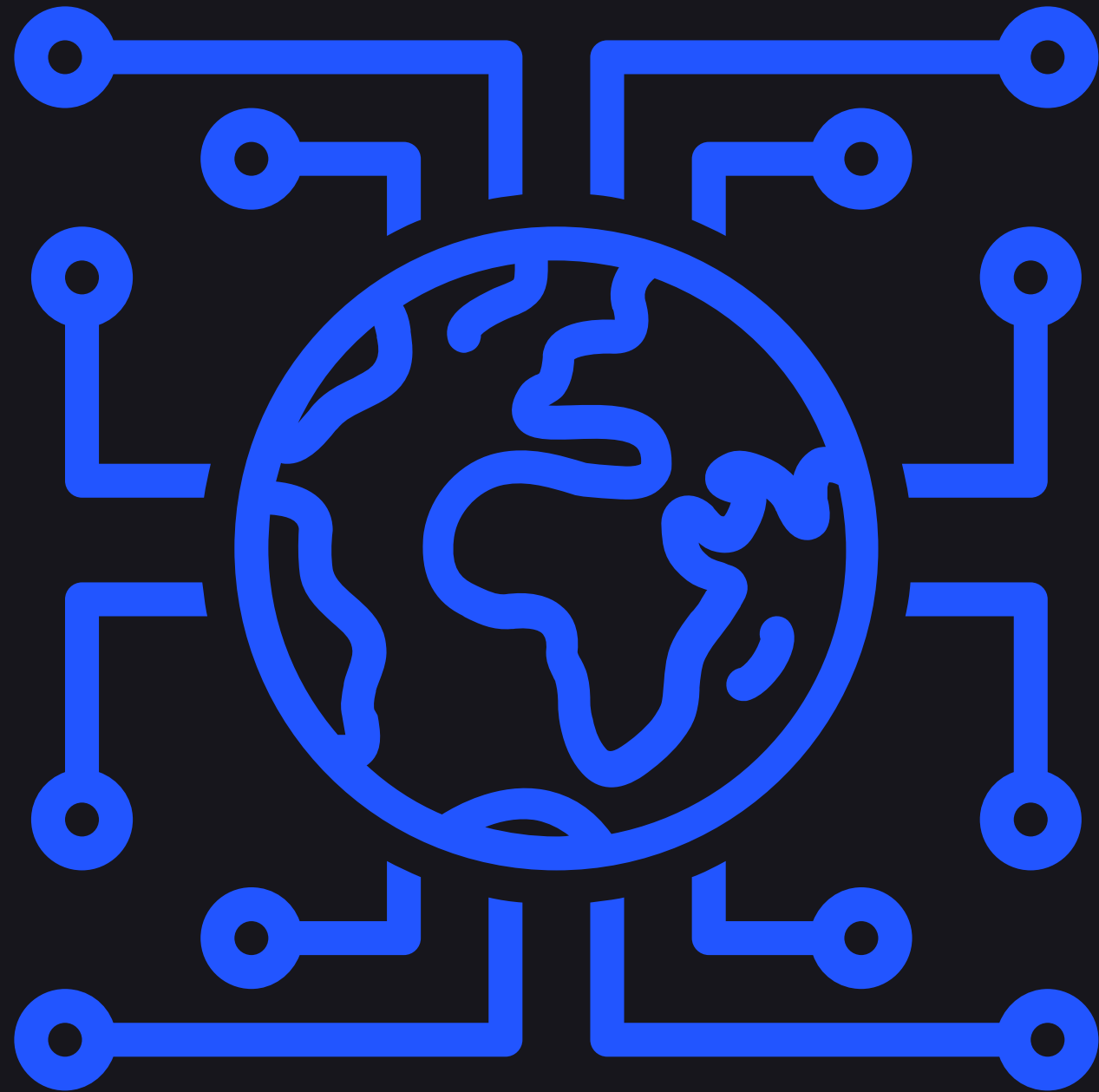


Publicação de Dados Abertos

- Preparação;
- Criação;
- Avaliação;
- Publicação.



Desafios de Dados na Web



- Metadados;
- Licença de Dados;
- Proveniência e qualidade dos Dados;
- Versionamento dos Dados;
- Identificação dos Dados;
- Formato dos Dados;
- Vocabulário de Dados;

Desafios de Dados na Web

- Acesso aos Dados;
- Preservação dos Dados;
- Enriquecimento dos Dados;
- Republicação dos Dados;



35 Boas Práticas para Dados na Web

Boa Prática 1: Fornecer metadados

Boa Prática 2: Fornecer metadados descritivos

Boa Prática 3: Fornecer metadados estruturais

Boa Prática 4: Fornecer informações sobre a licença de dados

Boa Prática 5: Fornecer informações de procedência dos dados

Boa Prática 6: Fornecer informações de qualidade de dados

35 Boas Práticas para Dados na Web

Boa Prática 7: Fornecer indicador de versão

Boa Prática 8: Fornecer o histórico de versão

Boa Prática 9: Usar URIs persistentes como
identificadores de conjuntos de dados

Boa Prática 10: Usar URIs persistentes como
identificadores dentro de conjuntos de dados

Boa Prática 11: Atribuir URIs para as versões dos
conjuntos de dados e séries

Boa Prática 12: Usar formatos de dados padronizados
legíveis por máquinas

35 Boas Práticas para Dados na Web

Boa Prática 13: Usar representações de dados que sejam independentes de localidade (locale neutral)

Boa Prática 14: Fornecer dados em formatos múltiplos

Boa Prática 15: Reutilizar vocabulários, dando preferência aos padronizados

Boa Prática 16: Escolher o nível de formalização adequado

Boa Prática 17: Fornecer download em massa (bulk download)

Boa Prática 18: Fornecer subconjuntos para conjuntos de dados extensos

35 Boas Práticas para Dados na Web

Boa Prática 19: Usar negociação de conteúdo para disponibilizar dados em formatos múltiplos

Boa Prática 20: Fornecer acesso em tempo real

Boa Prática 21: Fornecer dados atualizados

Boa Prática 22: Fornecer uma explicação para os dados que não estão disponíveis

Boa Prática 23: Disponibilizar dados por meio de uma API

Boa Prática 24: Usar padrões Web como base para construção de APIs

35 Boas Práticas para Dados na Web

Boa Prática 25: Fornecer documentação completa para as APIs

Boa Prática 26: Evitar alterações que afetem o funcionamento de sua API

Boa Prática 27: Preservar identificadores

Boa Prática 28: Avaliar a cobertura do conjunto de dados

Boa Prática 29: Coletar feedback de consumidores de dados

Boa Prática 30: Compartilhar o feedback disponível

Boa Prática 31: Enriquecer dados por meio da geração de novos dados

35 Boas Práticas para Dados na Web

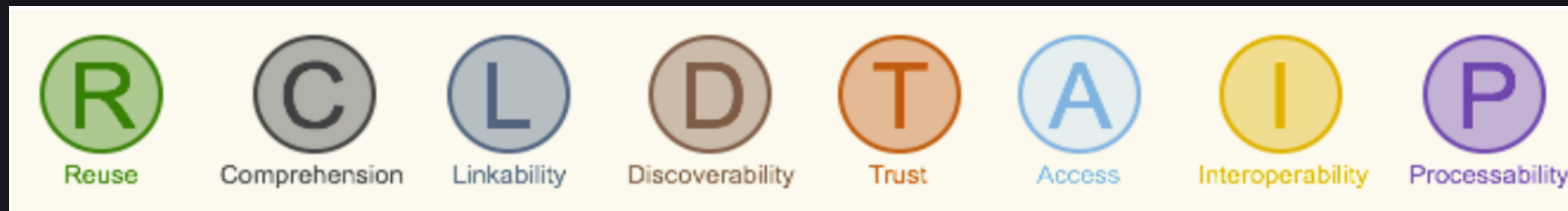
Boa Prática 32: Fornecer visualizações complementares

Boa Prática 33: Fornecer feedback para o publicador original

Boa Prática 34: Obedecer os termos de licença

Boa Prática 35: Citar a publicação original do conjunto de dados

Benefícios das Boas Práticas para Dados na Web



Reuse

- BP: Provide data license information
- BP: Provide versioning information
- BP: Provide version history
- BP: Use non-proprietary data formats
- BP: Provide data in multiple formats
- BP: Use a trusted serialization format for preserved data dumps
- BP: Enrich data by generating new metadata
- BP: Provide data provenance information
- BP: Provide data quality information
- BP: Use persistent URIs as identifiers

Trustworthy

- BP: Assess dataset coverage
- BP: Assign URIs to dataset versions and series
- BP: Provide data up to date
- BP: Update the status of identifiers
- BP: Gather feedback from data consumers
- BP: Provide information about feedback
- BP: Provide data provenance information
- BP: Provide data quality information

Comprehension

- BP: Provide metadata
- BP: Provide locale parameters metadata
- BP: Provide structural metadata
- BP: Provide descriptive metadata

Accessibility

- BP: Provide bulk download
- BP: Follow REST principles when designing APIs
- BP: Provide real-time access
- BP: Maintain separate versions for a data API
- BP: Assess dataset coverage

Linkability

- BP: Use persistent URIs as identifiers
- BP: Assign URIs to dataset versions and series

Discoverability

- BP: Provide descriptive metadata
- BP: Use persistent URIs as identifiers
- BP: Assign URIs to dataset versions and series

Processability

- BP: Use machine-readable standardized data formats
- BP: Enrich data by generating new metadata

Interoperability

- BP: Use standardized terms
- BP: Re-use vocabularies

Publicações do Ceweb.br



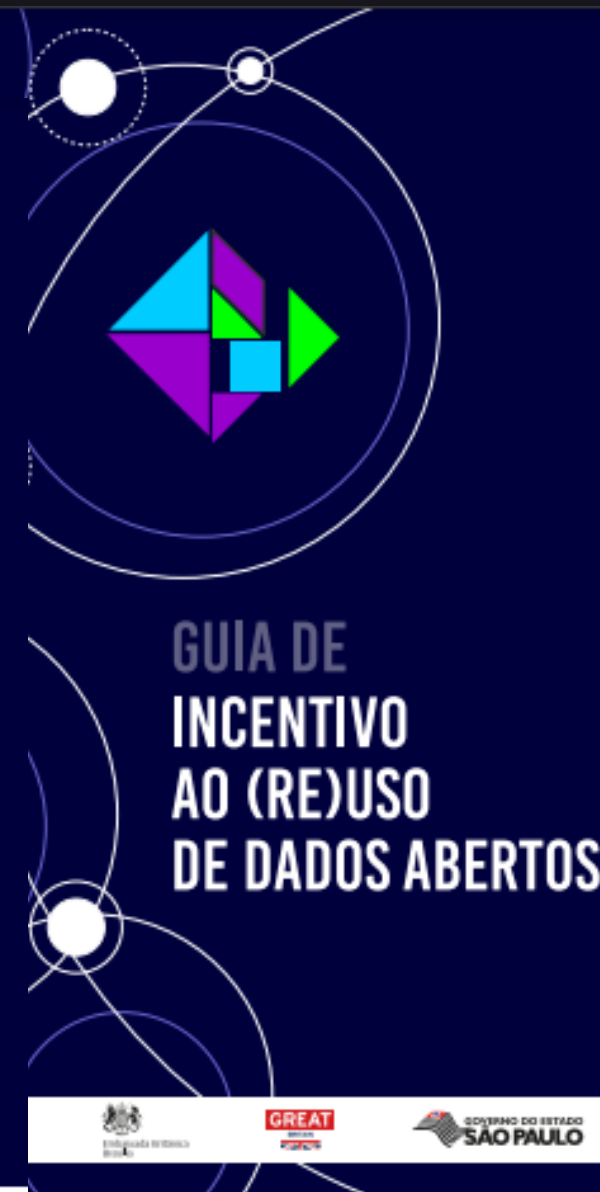
Acesse:

<https://acervo.ceweb.br>



<https://acervo.ceweb.br/acervos/itens/0f760d5c-01ee-49cd-b973-6e8128b81ac8>

Publicações do Ceweb.br



Acesse:
<https://acervo.ceweb.br>

<https://acervo.ceweb.br/acervos/itens/0f760d5c-01ee-49cd-b973-6e8128b81ac8>



Obrigada!

ceweb.br nic.br egi.br

Ana Eliza Duarte

email: anaeliza@nic.br

LinkedIn:

<https://www.linkedin.com/in/ana-duarte-a16196b5/>

www.ceweb.br