

Ficheiros & Metadados



Introdução e enquadramento



o. Introdução e enquadramento

o.1 Organização do trabalho

o.2 Algumas ideias sobre o assunto

0.2 Algumas ideias sobre o assunto - **Copyright** e **Propriedade Intelectual**

Artigo 27 da Declaração Universal dos Direitos Humanos

1. Todos os seres humanos têm o **direito de participar livremente na vida cultural** da comunidade, de **beneficiar** das **artes** e de participar no processo **científico** e dos seus benefícios.
2. Todos os seres humanos têm direito à **proteção dos interesses morais e materiais** decorrentes de qualquer produção científica, literária ou artística da qual seja autor.

Copiar? Partilhar? Arquivar?

0.2 Algumas ideias sobre o assunto - **Software Livre** e **Open Source**

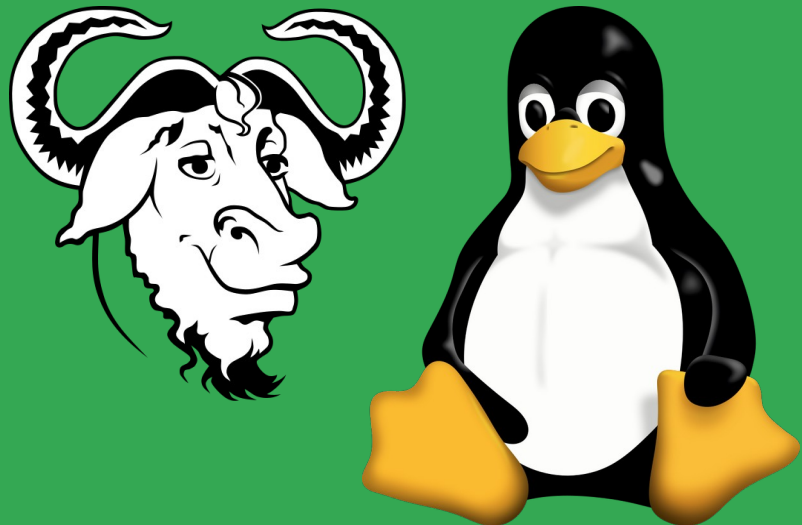


"Free software is a matter of **liberty, not price**. To understand the concept, you should think of free as in free speech, not as in free beer."

Richard Stallman



0.2 Algumas ideias sobre o assunto - **Software Livre** e **Open Source**



Liberdade 0 Executar o software para qualquer fim.

Liberdade 1 Estudar o seu funcionamento e alterá-lo.

Liberdade 2 Redistribuir cópias para que possa ajudar o próximo.

Liberdade 3 Modificar o programa e distribuir estas modificações, para benefício de toda a comunidade.

Software Livre (free software) não é software gratuito (freeware).

0.2 Algumas ideias sobre o assunto - **Software Livre** e **Open Source**



1983 GNU Project

1985 Free Software Foundation

1986 The Free Software Definition

1989 GNU General Public License

1991 Linux

1998 Open Source e OSI

2001 F/LOSS ou FLOSS

0.2 Algumas ideias sobre o assunto – **Código Aberto**

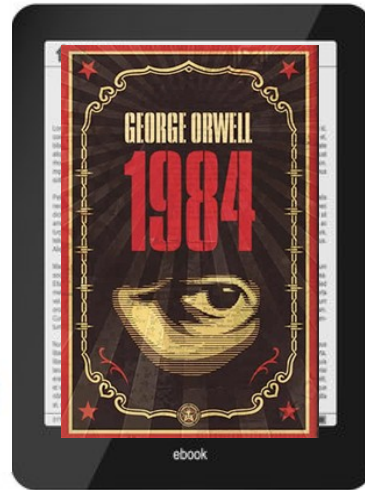


Open Source (Código Aberto) é software cujo código-fonte é disponibilizado publicamente, permitindo que qualquer pessoa possa ver, modificar e distribuir o seu código.

Open Source é Produto ou Metodologia?

O cisma Catedral vS Bazar

0.2 Algumas ideias sobre o assunto – **Código Aberto**

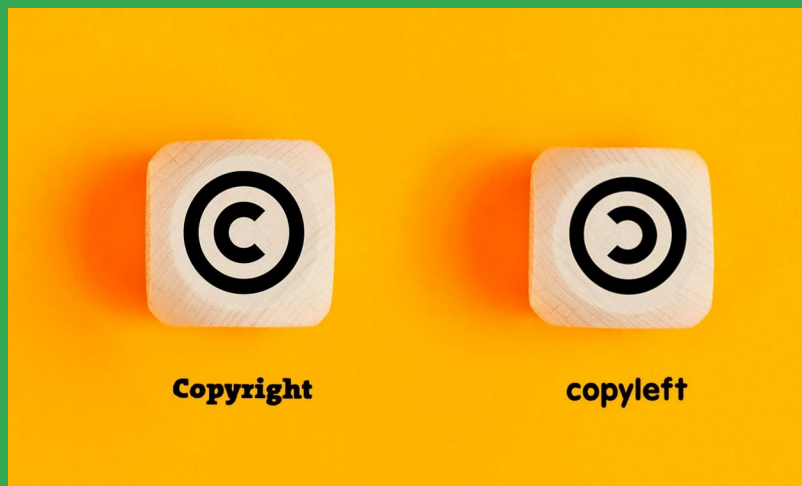


PHONE STORY

Phone Story is an educational game about the dark side of your favorite smart phone. Follow your phone's journey around the world and fight the market forces in a spiral of planned obsolescence.



0.2 Algumas ideias sobre o assunto - **Copyleft**



Copyleft é um conceito legal (e estratégia de licenciamento) baseado no copyright (direitos de autor) para garantir que uma obra e as suas derivadas permaneçam livres, assegurando as quatro liberdades do software livre ou os princípios da cultura livre.

Condição: Se redistribuir ou criar obra derivada tem de manter a mesma licença ou compatível. "bloqueia" a liberdade de tornar o livre em proprietário.

0.2 Algumas ideias sobre o assunto - **Formatos Abertos (livres)**

Regulamento Nacional de Interoperabilidade Digital

<https://www.acessibilidade.gov.pt/publicacao/rnid-versao-2018-regulamento-nacional-de-interoperabilidade-digital/>



Formato de ficheiro para armazenamento de dados digitais distribuído sob **licença livre/aberta**. É definido por uma especificação publicada abertamente, geralmente mantida por instituição pública ou internacional, e pode ser usado e implementado por qualquer pessoa para qualquer fim.

0.2 Algumas ideias sobre o assunto - **Acesso Aberto**



Princípios do Acesso Aberto:

- Gratuito para o leitor,
- Reutilização permitida,
- Disseminação rápida.

Acesso aberto não é só ler online gratuitamente.

Acesso Aberto (Open Access) é um movimento que defende a disponibilização gratuita e sem restrições de conteúdos académicos e científicos na Internet, permitindo que qualquer pessoa possa ler, descarregar, copiar, distribuir ou imprimir esses materiais sem barreiras financeiras ou legais.

“Open-access (OA) literature is digital, online, free of charge, and free of most copyright and licensing restrictions.”

Peter Suber

0.2 Algumas ideias sobre o assunto - **Acesso Aberto**



Inicialmente literatura científica, abriu o seu âmbito incluindo outros conteúdos.

1991 arXiv

1990s Open Archives Initiative

2001 An Open Letter to Scientific Publishers

2002 Budapest Open Access Initiative

2003 Bethesda Statement e Berlin Declaration, Public Library of Science

2008 RCAAP

2013 Aaron Swartz

2014 EU Horizon 2020 requer Open Access

0.2 Algumas ideias sobre o assunto - **Acesso Aberto**

Mais visível: Revistas OA (gold) & Arquivos/Repositórios OA (green)

Diamante (platina): textos académicos publicadas, distribuídas e preservadas sem custos para leitor ou autor.

Ouro: editor disponibiliza artigo (e conteúdo relacionado) imediatamente. Geralmente, existem APC (article processing charge) e artigos são publicados com licença CC que permite partilha e reutilização.

Verde: Auto-arquivamento pelo autor num repositório aberto, sem custos para leitor e autor.

Híbrido: muitos modelos de negócio... subscrição, embargo, leitor-pagador, etc.

Bronze: leitura gratuita no website, sem licença clara.

Black: Sci-Hub, Annas-archive.org

0.2 Algumas ideias sobre o assunto - **Cultura Livre**



Cultura Livre é um movimento e um conceito que defende a liberdade de partilha, modificação e distribuição de obras culturais (como música, filmes, livros, arte e conhecimento em geral), seguindo princípios semelhantes aos do software livre.

Esta obra está protegida pela lei. Não pode ser reproduzida, no todo ou em parte, qualquer que seja o modo utilizado, incluindo fotocópia e xerocópia, sem prévia autorização do Editor. Qualquer transgressão à lei dos Direitos de Autor será passível de procedimento judicial.

Todos os direitos reservados vs **Alguns** direitos reservados

Copyright (Inglês e EUA) integra **Direito de Autor** e **Direitos Conexos** (intérpretes ou executantes, produtores, editores)

direitos morais (autoria, integridade, ter o nome associado à obra) são direitos inalienáveis, irrenunciáveis e que não prescrevem.

direitos patrimoniais (exploração económica, inclui direito de publicar, distribuir ou reproduzir) são alienáveis, renunciáveis e prescrevem.

Direitos de autor no mundo digital

não existem custos
adicionais
associados a
consumidores
adicionais

consumo
não-rival

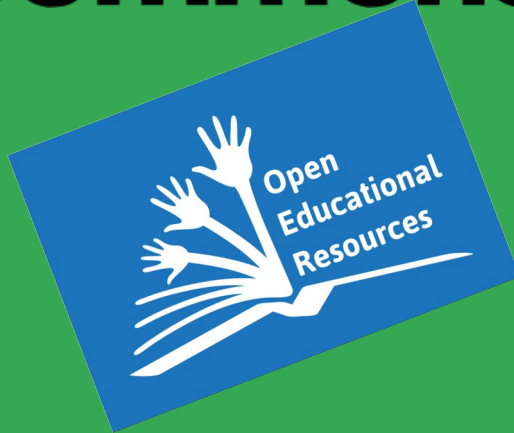


**bens digitais
tendem para a
abundância**

Preço
Copyright
EULA
mecanismos
de exclusão

criar
escassez
= criar
valor de
mercado

0.2 Algumas ideias sobre o assunto - **Cultura Livre**



1998 Sonny Bono Copyright Term Extension Act

2001 Creative Commons, Wikipedia

2002 Licenças Creative Commons, MIT OpenCourseWare, OER pela UNESCO

2003 Free Culture, Wikimedia

2004 Free Culture (livro)

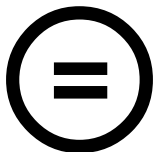
2006 Definition of Free Cultural Works, Elephants Dream

2009 Casa Branca, Wikipedia, Sita Sings the Blues

0.2 Algumas ideias sobre o assunto – **Licenças Creative Commons**



BY Atribuição obrigatória - O crédito deve ser dado ao autor/criador.



ND (Sem Derivações) Modificação não permitida - Não são permitidos derivados ou adaptações do trabalho



SA (Compartilha igual) Distribuído nos mesmos termos - As adaptações devem ser compartilhadas nos mesmos termos.



NC (não comercial) Uso comercial não permitido - Só o uso não comercial do trabalho é permitido



CC0 1.0 DEED

CC0 1.0 Universal



CC BY 4.0 DEED

Atribuição 4.0 Internacional



CC BY-SA 4.0 DEED

Atribuição-Compartilhalqual 4.0 Internacional



CC BY-NC 4.0 DEED

Atribuição-NãoComercial 4.0 Internacional



CC BY-NC-SA 4.0 DEED

Atribuição-NãoComercial-Compartilhalqual 4.0 Internacional



CC BY-ND 4.0 DEED

Atribuição-SemDerivações 4.0 Internacional



CC BY-NC-ND 4.0 DEED

Atribuição-NãoComercial-SemDerivações 4.0 Internacional

CREATIVE COMMONS LICENSES



**COPY
& PUBLISH**



**ATTRIBUTION
REQUIRED**



**COMMERCIAL
USE**



**MODIFY
& ADAPT**



**CHANGE
LICENSE**



PUBLIC DOMAIN



CC BY



CC BY-SA



CC BY-ND



CC BY-NC

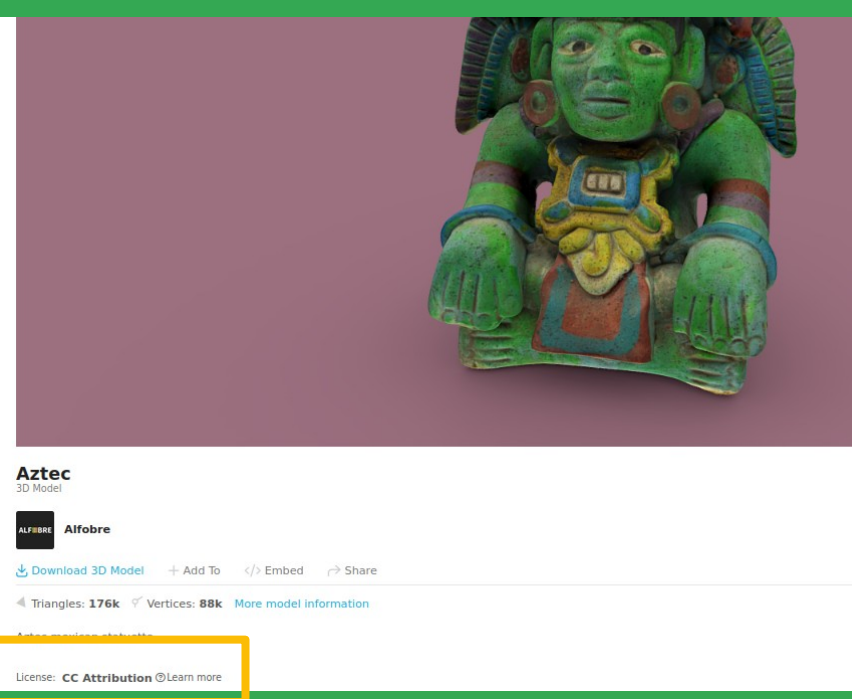
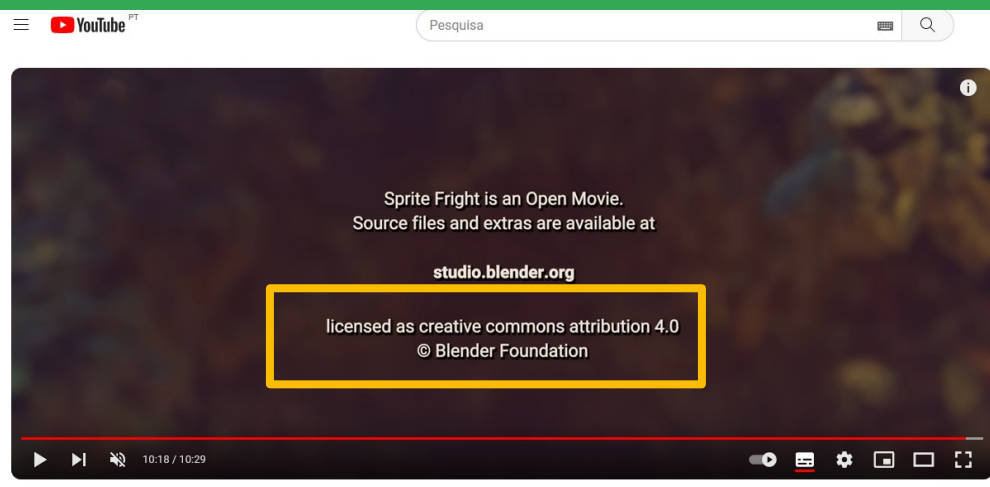


CC BY-NC-SA




CC BY-NC-ND





© European Union, 2023



The reuse policy of European Commission documents is implemented by Commission Decision 2011/833/EU of 12 December 2011 on the reuse of Commission documents (OJ L 330, 14.12.2011, p. 39). Unless otherwise noted, the reuse of this document is authorised under a Creative Commons Attribution 4.0 International (CC BY 4.0) licence (<https://creativecommons.org/licenses/by/4.0/>). This means that reuse is allowed provided appropriate credit is given and any changes are indicated.

0.2 Algumas ideias sobre o assunto – **Licenças Creative Commons**

Vamos jogar!

0.2 Algumas ideias sobre o assunto - **Dados Abertos**

OPEN DATA

Formatos Abertos (não proprietários), Licenças Livres e Metadados Descritivos

Dados abertos não é só pesquisar/consultar gratuitamente numa base de dados online.

Dados abertos são dados que podem ser livremente utilizados, reutilizados e redistribuídos por qualquer pessoa – sujeitos apenas, no máximo, à exigência de atribuição e partilha nos mesmos termos.

Open Data Handbook

0.2 Algumas ideias sobre o assunto - **Dados Abertos**



OpenStreetMap

zenodo



2004 OpenStreetMap

2005 The Open Definition vo.1 (Open Knowledge Foundation)

2009 Open Government Initiative

2012 EU Open Data Portal e European Data Portal

2015 Zenodo

2016 Princípios FAIR (findability, accessibility, interoperability, and reusability)

2023 Horizon Europe Programme Guide, v4 (15-10-2023) 'as open as possible, as closed as necessary'.

0.2 Algumas ideias sobre o assunto - **Princípios FAIR**

Findable (Localizável): dados devem ser facilmente encontrados por humanos e máquinas. Devem ter metadados claros e identificadores únicos.

Accessible (Acessível): dados devem estar disponíveis com protocolos padrão e livres. Metadados devem permanecer acessíveis mesmo se os dados forem removidos.

Interoperable (Interoperável): usar formatos e vocabulários universais. Metadados devem seguir padrões reconhecidos.

Reusable (Reutilizável): licenças claras e metadados devem explicar como, quando e por quem os dados foram gerados.

0.2 Algumas ideias sobre o assunto - **Ciência Aberta**



Ciência Aberta (Open Science) é um movimento que visa tornar a investigação científica mais transparente, acessível e colaborativa, removendo barreiras ao conhecimento.

Pilares: Publicações em Acesso Aberto, Dados de Investigação Abertos, Revisão por Pares Aberta, Ciência Cidadã, Software e Ferramentas Livres, Educação Aberta.

0.2 Algumas ideias sobre o assunto - **Ciência Aberta**



1985 Open Science and Closed Science: Tradeoffs in a Democracy

2014-2020 Horizon 2020

2016 Grupo de trabalho Política Nacional Ciência Aberta

2021 UNESCO Recommendation on Open Science

2021-2027 Horizon Europe

0.2 Algumas ideias sobre o assunto - **Metadados**

Metadados são **dados sobre dados** – informações estruturadas que descrevem, explicam ou ajudam a organizar outros dados, tornando-os mais fáceis de encontrar, gerir e interpretar.

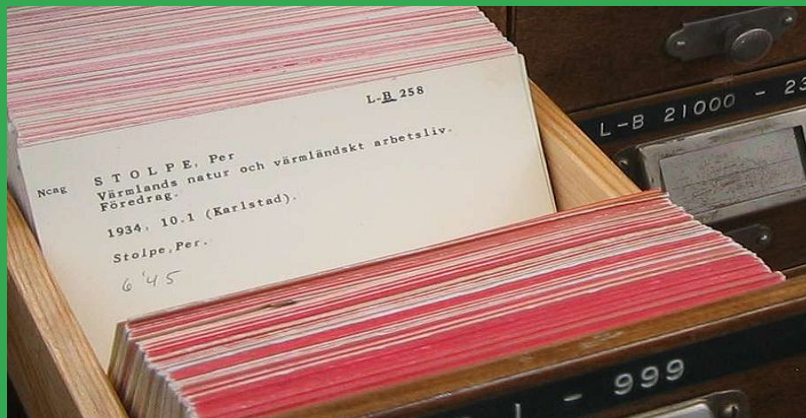
Ficheiro de música: Dados: O ficheiro MP3 em si.
Metadados: Título, artista, álbum, duração, género, etc.

Fotografia digital: Dados: A imagem (píxeis).
Metadados: Data, hora, localização (GPS), modelo da câmara, etc.

Artigo científico: Dados: O texto do artigo.
Metadados: Autor, resumo, palavras-chave, DOI, instituição, etc.

<https://iem.webgis1.com/catalogue/#/dataset/397>

0.2 Algumas ideias sobre o assunto - **Metadados**



EXCLUSIVO SEGURANÇA

Conselho que fiscaliza secretas aponta aumento de ameaças e pede acesso a metadados

Conselho de fiscalização dos serviços de informações pede revisão constitucional e diz que as secretas precisam de mais meios humanos e têm dificuldade em reter profissionais qualificados.

Metadados existiam antes dos digital.

No mundo digital, tudo tem metadados. Metadados existem em todos os ficheiros digitais.

Metadados são criados mesmo quando não sabemos.

0.2 Algumas ideias sobre o assunto - **Metadados**

1. **Smartphones**: Metadados EXIF nas fotos guardam localização GPS e hora da captura (modelo da câmara, exposição, etc.). Google Fotos usa para criar álbuns automáticos ("Viagens de 2023").
2. **Cartões de Transporte** (Metro): guardam horários de uso, estações frequentes e saldo. Planeamento e gestão de transportes usa esses dados para planear rotas de autocarros.
3. **Likes e Partilhas**: Cada like/share tem metadados como hora, dispositivo usado ou localização aproximada. O Instagram sugere "Amigos que talvez conheças" cruzando esses dados.
4. **Faturas Eletrónicas**: Para além do valor, têm metadados como NIF do vendedor, tipo de transação, etc. A Autoridade Tributária usa-os para detetar fraudes.

0.2 Algumas ideias sobre o assunto - **Metadados**

5. **Spotify**: Metadados de músicas incluem batidas por minuto (BPM), emoção associada (ex.: "energético"), etc. Spotify usa metadados para recomendações.
6. **E-mails**: IP do remetente, servidores por onde passou e horários.
7. **Livros Digitais**: Metadados como tempo de leitura, trechos sublinhados e páginas mais abandonadas. A Amazon sugere livros com base no ritmo de leitura.
8. **Internet**: pesquisa web (crawlers), visitas, metatags, etc.

```
<meta charset="UTF-8">
<meta name="viewport" content="width=device-width, initial-
scale=1">
<meta name="description" content="A brief description of your
page">
<meta name="keywords" content="keyword1, keyword2, keyword3">
<meta name="author" content="Your Name">
<meta name="robots" content="index, follow">
<meta property="og:title" content="Your Page Title">
```




Se os ficheiros são os
tijolos...

Os metadados são a
argamassa.

