

GUIA DADOS ABERTOS

AMA | dados.gov

2016



Índice

1. O que são Dados Abertos?	3
1.1. <i>Open Government</i> e informação pública.....	3
1.2. Dados abertos: uma definição	5
1.3. Benefícios dos “Dados Abertos”	7
1.4. Questões de privacidade	8
2. O potencial dos Dados Abertos	9
2.1. Transparência e cidadania.....	9
2.2. Potencial económico.....	13
2.3. Outros tipos de utilização dos Dados Abertos.....	15
3. Como abrir os dados.....	18
3.1. Planeamento	18
3.2. Escolha dos dados a abrir.....	20
3.3. Tipos de utilizadores	21
3.4. Dados prioritários	23
3.5. Disponibilização dos dados	24
3.6. Estratégia de divulgação.....	26
3.7. Manutenção dos dados	28
3.8. Exemplos de iniciativas de dados públicos abertos.....	29
3.9. Dados.Gov: O portal nacional de dados abertos	31
4. Reutilizar os dados.....	33
4.1. Formatos abertos.....	34
4.2. Estrutura de dados.....	37
4.3. Metadados	39
4.4. Licenças abertas.....	42
4.5. Escolha da licença	43
4.6. Publicar a licença.....	44

1. O QUE SÃO DADOS ABERTOS?

1.1. *Open Government* e informação pública

A definição inicial de *Open Government*, ou “Governo Aberto”, é a de um Governo cujas principais características são, segundo a OCDE¹ a transparência, a participação e a colaboração.

Nesse sentido, Governo Aberto significa que Governo e Administração Pública estão abertos ao escrutínio e à contribuição dos cidadãos, da sociedade civil e de organizações várias, para que haja um maior envolvimento por parte de todos e criação de valor público.

A demanda por um Estado mais transparente, e que fomente a participação cívica, não é nova, sendo que nos últimos anos as questões da transparência e da cidadania têm sido impulsionadas pelo desenvolvimento das novas tecnologias. Estas questões, passaram a estar enquadradas nos movimentos Open, e ganharam um importante impulso político em todo o mundo a partir de 2008, quando vários governos, entre os quais os Estados Unidos da América, o Canadá, o Reino Unido e a Nova Zelândia, começaram a divulgar e a implementar medidas destinadas a disponibilizar a sua informação pública².

À escala global, o aparecimento de iniciativas como a *Open Government Partnership*³, criada em setembro de 2011, e cujos principais objetivos são promover a transparência, lutar contra a corrupção e dar mais poder aos cidadãos, veio dar mais força ao movimento de abertura dos Governos. Esta iniciativa agrega um conjunto de países que ao aderir se comprometem a atingir metas na área da abertura e transparência.

Informação do Sector Público

Uma das preocupações transversais à maioria dos países relaciona-se com a importância da disponibilização de informação do sector público. A facilidade com que os cidadãos e sociedade em geral podem aceder aos dados produzidos pelos Governos e Administrações Públicas.

A questão do acesso à informação do sector público também ganha nova relevância à luz dos avanços tecnológicos – muito associados à denominada *web 2.0* - que permitem aceder a informação de forma mais direta, prática e, sobretudo, conseguir reutilizar essa informação

¹ “Building an open and innovative Government for better policies and service delivery.” Paris: OECD, 8-9 June 2010

² Manual de Dados Abertos

³ <http://www.opengovpartnership.org/>

para diversos fins ou contextos. A emergência de plataformas e modelos colaborativos, o crescimento dos movimentos *Open* (*open source*, *open software*, etc.) reforçou as condições para uma maior participação dos cidadãos na sociedade. Nesse sentido, os Governos foram obrigados a repensar e redefinir o seu papel como produtores e fornecedores de informação, disponibilizando o maior número de dados públicos e facilitando, incentivando e permitindo que a sociedade civil tenha acesso facilitado aos mesmos

No quadro da União Europeia, a questão da reutilização de informação do sector público deu um impulso decisivo em 2013, com a revisão da Diretiva Europeia da Informação do Setor Público (Diretiva 2003/98/EC revista pela Diretiva 2013/37/EU⁴).

Também conhecida como Diretiva PSI (Public Sector Information), esta diretiva europeia passou a equiparar o potencial económico da informação detida pelo setor público ao seu valor de transparência, encorajando os países europeus a colocarem mais dados públicos em formatos abertos à disposição da sociedade civil para reutilização, fomentando o desenvolvimento de novos produtos e serviços baseados em dados abertos.

Em Portugal, é a Lei de Acesso aos Documentos Administrativos⁵ (LADA) que regula o acesso aos dados públicos. Como diz o primeiro artigo, intitulado Administração Aberta: “O acesso dos cidadãos aos documentos administrativos é assegurado pela Administração Pública de acordo com os princípios da publicidade, da transparência, da igualdade, da justiça e da imparcialidade.”

A enorme quantidade de dados que é gerada e centralizada pelo Estado congrega em si um enorme potencial de utilização e de desenvolvimentos que podem ser úteis e importantes tanto para o Estado como para a sociedade civil e mundo empresarial. A grande maioria desses dados já é, por lei, considerados públicos. O grande desafio (e a maior preocupação do movimento dos Dados Abertos) passa por facilitar o seu acesso e reutilização, beneficiando vários grupos e sectores da sociedade:

- Os cidadãos, que passam a ter um acesso mais imediato a informação que lhes pertence por direito, reforçando a visão de transparência e prestação de contas do Estado perante os eleitores;
- As instituições governamentais, que se tornam mais transparentes e têm a hipótese de se tornarem mais eficientes e eficazes, reforçando também o seu papel de serviço público;

⁴ <https://ec.europa.eu/digital-single-market/en/european-legislation-reuse-public-sector-information>

⁵ <http://www.cada.pt/modules/smartsection/item.php?itemid=41>

- O sector empresarial, que pode reutilizar informação pública para criar aplicações, plataformas ou serviços com elevado potencial comercial;

Entre muitos outros, incluindo sectores como jornalismo, investigação universitária ou mesmo organizações não-governamentais com preocupações cívicas. Ao longo deste guia, à medida que aprofundarmos os processos e metodologias de abertura e reutilização de *open data*, iremos também fornecer vários de exemplos de utilização de dados.



Share-PSI 2.0

Boa Prática: Disponibilizar Informação do Sector Público a custo zero

Estabelece como boa prática a disponibilização de informação pública sem qualquer tipo de custo, marginal ou outro, para a sua reutilização.

Provide PSI at zero charge, Disponível em: <https://www.w3.org/2013/p-psi/bp/zero/>

1.2 Dados abertos: uma definição

Recorrendo à definição do site *Open Definition*⁶, dados abertos são “dados que podem ser livremente utilizados, reutilizados e redistribuídos por qualquer um.”

Dito de outra forma, dados abertos são dados que podem ser utilizados sem qualquer tipo de restrição. No âmbito deste guia estamos a considerar dados abertos como sendo “dados abertos públicos”, embora também existam dados abertos que não sejam provenientes do sector público.

Em relação aos dados públicos, é frequente o Estado impor alguns limites àquilo que os cidadãos (e as empresas) podem fazer com a informação pública. Os mais comuns são vedar a reutilização para fins comerciais, através da imposição de algum tipo de restrição sobre o seu licenciamento, permitindo apenas a exploração académica ou a título individual. Ou seja, mesmo que os dados estejam disponíveis para serem descarregados, o cidadão não os pode reutilizar sem limitação.

Este é um dos pontos mais prementes do movimento pelos dados abertos – a abolição de qualquer tipo de limitação à sua reutilização comercial. A ideia de que qualquer pessoa ou entidade pode utilizar, transformar ou adaptar os dados públicos para rentabilizar num

⁶ <http://opendefinition.org/>

contexto de negócio. A título de exemplo, o termo “dados abertos”, quando aplicado ao sector público (como acontecerá ao longo deste guia), pode referir-se à informação proveniente de vários setores da Administração Pública e relativa à sua atuação, como por exemplo: contratação pública, emprego público, estatísticas económicas e financeiras, despesas e receitas públicas, resultados eleitorais, georreferenciação de moradas e serviços públicos, horários de transportes, indicadores de qualidade de serviço, entre muitos outros. Grande parte desta informação pode até já estar disponível publicamente, mas para que um determinado conjunto de dados governamentais possa ser classificado como Aberto é necessário que não existam restrições ao seu acesso, sejam estas legais ou políticas, tecnológicas ou financeiras.

Assim, distingue-se “dados abertos” de dados que sejam apenas disponibilizados ao público. Os dados abertos são sempre abrangidos por licenças “abertas”, que permitem reutilização comercial. Se assim não for – não poderão ser considerados “dados abertos”. A temática das licenças é também abordada no capítulo 4.

As 3 Leis de Dados Abertos

David Eaves, professor na Harvard Kennedy School, nos Estados Unidos da América, e defensor de uma política de transparência em termos de dados governamentais resume desta forma as 3 leis dos Dados Abertos:

1. “Se os dados não podem ser encontrados ou indexados na *Web*, eles não existem.”
2. “Se os dados não estão disponíveis num formato aberto e legível por máquina, eles não podem ser reutilizados.”
3. “Se dispositivos legais não permitem que eles sejam partilhados, eles não são úteis.”

A clareza com que se define o conceito de Dados Abertos e a existência de regras ou leis como as descritas acima é o que permite que exista uma certa uniformização na forma como a informação é disponibilizada e, conseqüentemente, o cruzamento da informação.

Para isso, de uma forma resumida devem verificar-se as seguintes condições:

1. **A garantia de acessibilidade:**
 - É fácil encontrar estes dados através de uma pesquisa simples na Internet.
 - Sem custos para o utilizador.

- A informação está completa, é clara e está disponível em ficheiros de uso comum que possam ser processados automaticamente por *softwares* e sistemas, possibilitando a criação de *apps*, sites ou plataformas por qualquer pessoa ou entidade.

2. A utilização também é “aberta”:

- Permitir que os conteúdos sejam livremente utilizados, podendo ser ainda reutilizados e redistribuídos.
- Possibilidade de cruzar estes dados com outras fontes de informação e gerar uma combinação de bases de dados abertas.
- Disponibilização dos dados tal como colhidos da fonte, sem agregação ou modificação, e o mais rápido possível, quanto mais recentes e atuais, mais úteis.

3. A ausência de restrições à reutilização:

- Estão disponíveis para qualquer pessoa, sem necessidade de registo ou qualquer outro procedimento que impeça ou monitorize o acesso.
- Qualquer indivíduo ou coletividade pode aceder e utilizar a informação, não havendo restrições legais associadas por exemplo ao setor de atividade ou esfera de atuação (académica, governamental, comercial, etc.), nem ao objetivo da consulta e eventual utilização.

1.3. Benefícios dos “Dados Abertos”

A transparência e o escrutínio democrático, assim como ajudar a uma maior participação dos cidadãos na sociedade, são benefícios indiretos e alcançáveis a médio ou longo prazo.

O benefício mais prático e imediato é, sem dúvida, a transparência, o Estado passa a operar numa cultura em que a sua informação está acessível muito rapidamente e de forma mais clara para sociedade em geral e para próprio Estado. E isto materializa um impacto profundo e de extrema importância.

Este benefício central é secundado pelos benefícios que traz a interligação de diferentes conteúdos por diferentes agentes da sociedade, gerando novas e melhores soluções de disponibilização da informação para os cidadãos. É essa interligação que permite também explorar o potencial existente na partilha e cruzamento de dados oficiais, passando estes a ter utilizações muito mais abrangentes, úteis, relevantes e geradoras de valor para a sociedade.

Para um Governo que pratique uma política de dados abertos, há também muitos benefícios. Desde logo há um novo conhecimento a partir da combinação de bases de dados e padrões, que poderá levar à introdução de inovações, decorrentes da análise do impacto das suas políticas, e a melhorias na eficiência dos seus serviços.

1.4. Questões de privacidade

Nem toda a informação do sector público deve ser tornada pública. Há um vasto conjunto de dados na Administração Pública que deve continuar na esfera de acesso restrito, seja por razões de segurança, razões legais, ou direito à privacidade dos cidadãos. As políticas de dados abertos não estipulam que se “abra” toda a informação do Estado, apenas aquela que pode ser considerada pública.

As considerações relativamente à abertura de dados de sectores específicos cabem às entidades que os gerem em consonância com organismos como a Comissão Nacional de Proteção de Dados.

Quando se fala de dados abertos fala-se sobretudo em dados governamentais que já estão ou deveriam estar disponíveis para a sociedade e que, assim, têm o potencial de se tornarem abertos, garantindo sua reutilização em novos projetos.

2. O POTENCIAL DOS DADOS ABERTOS

2.1. Transparência e cidadania

‘Todos os seres humanos têm direito à liberdade de opinião e de expressão; esse direito inclui a liberdade de ter opiniões sem sofrer interferência e de procurar, receber e divulgar informações e ideias por quaisquer meios, sem limite de fronteiras’.

A frase acima faz parte do artigo 19, relativo ao direito à informação, da Declaração Universal dos Direitos Humanos, adotada pela Organização das Nações Unidas (ONU) em 1948. Em Portugal, após sua revisão de 1982, a Constituição da República Portuguesa passou também a referir o direito à informação.

Há um claro benefício ao nível individual que advém desse direito humano fundamental à informação e à liberdade. Mas, em última análise, ao permitir que qualquer cidadão tenha acesso facilitado e livre aos dados produzidos pelo Governo, é a sociedade como um todo que beneficia.

Nas últimas décadas, o rápido desenvolvimento das Tecnologias de Informação e Comunicação e a democratização no acesso a plataformas e estruturas tecnológicas, alteraram radicalmente os paradigmas que criação e transformação de informação. O potencial da reutilização de dados abertos é assim indissociável da capacidade de trabalhar esses dados recorrendo a mecanismos automáticos, integrando-os em aplicações ou plataformas de acesso livre (através da *web*) ou disponíveis através de modelos económicos.

Para melhor ilustrar o potencial da reutilização de informação do sector público e sua relação com a transformação digital, passamos a elencar uma série de exemplos de plataformas ou aplicações que utilizam informação pública para criar diversos tipos de valor.

Exemplos de utilização dos dados abertos

TRANSPARÊNCIA



Portal de Transparência Municipal

Lançado em julho de 2014, o Portal de Transparência Municipal disponibiliza um conjunto de indicadores relativos à gestão dos 308 municípios portugueses e entidades intermunicipais.

<https://www.portalmunicipal.pt>

Exemplos de utilização dos dados abertos



Portal de Transparência do Serviço Nacional de Saúde

Iniciativa de dados abertos do Ministério da Saúde que disponibiliza dados relativos às operações e transações das diversas entidades que integram o Sistema Nacional de Saúde (SNS)

<https://transparencia.sns.gov.pt>



Índice de Transparência Municipal

Publicado anualmente desde 2013 pela Transparência e Integridade, Associação Cívica (TIAC), o Índice de Transparência Municipal mede o grau de transparência das Câmaras Municipais através de uma análise da informação disponibilizada aos cidadãos nos seus *web sites*, com base em 76 indicadores agrupados em sete dimensões.

<http://poderlocal.transparencia.pt/>

MONITORIZAÇÃO DA GOVERNAÇÃO E CONTAS PÚBLICAS

WHERE DOES MY MONEY GO?
Showing you where your taxes get spent



Para Onde Vão Os Meus Impostos (Where Does My Money Go?)

O serviço *online* “Para Onde Vão Os Meus Impostos” é um dos projetos mais emblemáticos da exploração dos dados abertos em prol da transparência. <http://app.wheredoesmymoneygo.org/>

Dollars for Docs: How Industry Dollars Reach Your Doctors

Uma ferramenta *online* que recorre aos dados públicos para monitorizar os pagamentos da indústria farmacêutica a médicos e unidades de saúde. <https://projects.propublica.org/docdollars/>



App #SICONV

A *app* “#SICONV”, lançada pelo Ministério do Planeamento do Governo do Brasil, permite a qualquer cidadão monitorar a execução de convénios firmados entre entidades de um município e a União, e verificar a prestação de contas relativa ao uso dos recursos e o número de emendas parlamentares propostas e assinadas. Os dados estão disponíveis em formato de gráficos, tabelas ou imagens.

<https://play.google.com/store/apps/details?id=com.siconv.mpog.android>



European Structural & Investment Funds Data

Desenvolvida pela Comissão Europeia, esta ferramenta *online* permite explorar os dados da execução dos 531 programas de financiamento europeus à data de novembro de 2015. Os dados podem ser explorados por tema, país ou por programa de financiamento, permitindo perspetivas sobre o investimento planeado agregado por Estado membro, por tema ou por programa de financiamento.

<https://cohesiondata.ec.europa.eu/>

Exemplos de utilização dos dados abertos

ACESSO A SERVIÇOS PÚBLICOS



Mapa do Cidadão

Lançada pela Agência para a Modernização Administrativa, a aplicação Mapa do Cidadão fornece informação sobre os serviços públicos à disposição dos cidadãos. Sobre o mapa, o utilizador tem a possibilidade de saber, por exemplo, quais os serviços públicos mais próximos, que serviços estão disponíveis num balcão, qual o contacto e o horário de funcionamento do serviço e interagir com o sistema de atendimento usado nas Lojas do Cidadão e nos Espaços do Cidadão. A plataforma baseia-se nos dados disponíveis no portal nacional de dados abertos Dados.Gov.pt e está disponível na *Web*, nos *smartphones* e nos *tablets*. <https://www.mapadocidadao.pt/>

DEMOCRACIA



Congresoscopio

O projeto “Congresoscopio” desenvolvido pelo jornal argentino La Nación explora os dados da atividade legislativa do país para responder à questão “Cómo votan nuestros legisladores?”. <http://votaciones.lanacion.com.ar>



Declaraciones Juradas

Através da aplicação *online* “Declaraciones Juradas”, o La Nación dá aos seus leitores a possibilidade de ficarem a conhecer em detalhe o património dos funcionários públicos da Argentina. <http://interactivos.lanacion.com.ar/declaraciones-juradas>

VIDA URBANA



Lisboa Aberta

A plataforma “Lisboa Aberta”, lançado no início de 2016 pela Câmara Municipal de Lisboa, oferece um portal de dados abertos que disponibiliza informação em vários domínios da vida da capital portuguesa, incluindo transportes, serviços, cultura ou dados estatísticos.

<http://dados.cm-lisboa.pt/>



OpenGrid

Os habitantes da cidade norteamericana de Chicago dispõem de um portal, OpenGrid, que dá possibilidade usar um computador ou dispositivo móvel para explorar os dados da cidade, descobrir empresas prestes a abrir, edifícios novos ou renovados, serviços da cidade e pontos de interesse, entre muitos outros aspetos.

<http://opengrid.io/>

Exemplos de utilização dos dados abertos



Urban Data Portal

A UN-Habitat, a agência especializada da ONU que se dedica à promoção da sustentabilidade social e ambiental das cidades, lançou o portal de dados abertos Urban Data Portal que disponibiliza dados sobre as tendências que mais têm vindo a pressionar o ambiente urbano nas últimas décadas.

<http://urbandata.unhabitat.org/>

TRÂNSITO



Tráfego Madrid

Com base nos dados disponíveis no portal de dados abertos do Ayuntamiento de Madrid, foi desenvolvida a aplicação “Tráfego Madrid” que permite aos condutores da capital espanhola acompanhar em tempo real os incidentes de trânsito, obras e semáforos de controlo de passadeiras, entre outros aspetos.

<http://informo.munimadrid.es/>

SAÚDE



GeoHealth

A *app* “GeoHealth” foi desenvolvida numa competição de programação de aplicações Hackathon HeartBits, que teve lugar no Minho. Esta ferramenta dá a possibilidade de analisar os fatores de risco para a saúde com base na localização geográfica do utilizador.

<http://seium.org/heartbits/>



A Tu Servicio

A ferramenta “A Tu Servicio” permite comparar os prestadores de saúde do Uruguai em diversos indicadores, incluindo tempos de espera e recursos humanos disponíveis, entre outros. Lançada pelo portal nacional de dados abertos e Ministério de Saúde Pública do país, esta aplicação visa promover uma tomada de decisão informada dos utentes e fomentar a transparência do sistema.

<http://atuservicio.uy/>



Temporada de Pases

Também no Uruguai foi desenvolvida uma aplicação que recorre aos dados abertos para permitir a um utente saber qual o prestador de saúde mais adequado à sua situação. Denominada “Temporada de Pases”, a ferramenta *online* foi desenvolvida pela organização da sociedade civil “180 CIENCIA y DATA” com base nos dados abertos do Ministério da Saúde Pública do Uruguai.

<http://mutualistas.datauy.org/>

Exemplos de utilização dos dados abertos

EDUCAÇÃO

**Mejora Tu Escuela**

O portal mexicano “MejoraTuEscuela.org” oferece uma plataforma de participação cidadã que abre os dados sobre as instituições de ensino do país para convidar à participação dos cidadãos, com destaque para os encarregados de educação. Esta plataforma *online* independente e sem fins lucrativos permite pesquisar instituições de ensino, saber detalhes da oferta e qualidade de serviço de cada escola e comparar escolas. Nesta plataforma, os encarregados de educação podem ainda apresentar sugestões de melhoria a introduzir nos estabelecimentos de ensino. <http://www.mejoratuescuela.org/>

**Ualojamiento**

Ajudar os alunos que têm necessidade de encontrar casa para morar durante o seu percurso universitário é o objetivo da *app* “Ualojamiento”, desenvolvida pela Universidad de Alicante com base nos *datasets* disponibilizados no portal de dados abertos da universidade. Disponível nas plataformas iOS, Android e Windows Phone, a *app* “Ualojamiento” agrega a oferta de alojamentos à disposição dos alunos de Alicante e fornece os dados essenciais de cada casa disponível para aluguer, incluindo o tipo de casa, as características, o preço e a localização. A aplicação permite ainda visualizar as ofertas sobre o mapa. <http://datos.ua.es/>

2.2. Potencial económico

Para além dos ganhos na área da transparência e da cidadania, há um tremendo potencial económico associado à abertura dos dados gerados pela Administração Pública que se concretiza num estímulo direto à economia e à criação de novos postos de trabalho.

Um estudo⁷ económico levado a cabo no âmbito do Portal de Dados Europeu⁸ estima que o mercado direto de Dados Abertos europeu atinja, em 2020 os 75.7 mil milhões de euros e que o número de empregos criados na área, até esse ano, chegue quase aos 100 mil.

⁷ <http://www.europeandataportal.eu/en/node/101>

⁸ <http://www.europeandataportal.eu/>

Em termos práticos, os benefícios materiais de uma política de dados abertos variam de acordo com a utilização que deles será feita.

- **Benefícios para o setor público**

Verifica-se, em muitos países, que as Administrações Públicas que decidem disponibilizar os seus dados passam também a beneficiar diretamente da abertura dessa informação, muitas vezes localizada em silos ou em formatos de reduzida interoperabilidade.

A abertura de dados também permite analisar melhor o impacto da atividade (podendo medi-lo) e também a existência de obstáculos internos, essa reutilização dos dados gera muitas vezes um aumento da eficiência do organismo, por exemplo através da redução de carga de trabalho e/ou de custos administrativos.

A interligação e integração de diferentes bases internas traz potenciais ganhos para a gestão pública, quer em termos de qualidade (por exemplo evitando lapsos ou desatualizações decorrentes da duplicação de documentos), quer em termos de eficiência (por exemplo na redução de recursos humanos envolvidos).

- **Benefícios para o setor privado**

A abertura responsável e consistente de dados governamentais permite que qualquer empreendedor, ou grupo de empreendedores, use sua criatividade e visão ‘de fora’ para construir ferramentas, soluções e avanços tecnológicos que utilizem informação do sector público.

Estas aplicações não deverão ser entendidas como concorrentes da ação do Estado, mas sim uma forma de gerar valor económico e social a partir da, e para a, Administração Pública. Ao disponibilizar essa informação e estimular a sua reutilização, o Governo permite abrir a possibilidade do sector privado desenvolver novos produtos e serviços que poderão ser dirigidos para o mercado de consumo, o mercado empresarial ou o próprio sector público. Idealmente, dentro do próprio sector empresarial a abertura de dados poderá fomentar o surgimento de novas empresas com modelos de negócio onde a criatividade e a inovação são os pilares fundamentais.

Num contexto económico em que os dados são cada vez mais valiosos, existem claro benefícios que poderão surgir da interligação de dados públicos com dados privados, e o seu tratamento num contexto de *Big Data* ou de plataformas de *Business Intelligence*.



Figura 1 - Dados Abertos podem ajudar (European Data Portal)

Há inúmeros exemplos de empresas e serviços que surgiram a partir da abertura de dados, contribuindo para a geração de empregos e de riqueza, criando uma interdependência saudável entre o Governo e a iniciativa privada (ex. [The Weather Channel](#), [Moovit](#), [ReCollect](#)).

 **Share-PSI 2.0** **Boa Prática: Abrir dados de transportes públicos**

Um sistema de transportes públicos mais inteligente, ecológico e integrado é uma das grandes prioridades na área das infraestruturas e mobilidade a nível mundial. A abertura de dados relativos a transportes públicos (i.e. horários, dados em tempo real, paragens, acessibilidade, frotas, etc.) pode fomentar melhorias consideráveis a nível de experiência para os utilizadores e eficiência para os operadores.

Open Up Public Transport Data. Disponível em: <https://www.w3.org/2013/share-psi/bp/ptd/>

2.3 Outros tipos de utilização dos Dados Abertos

Os benefícios materiais decorrentes da abertura de dados governamentais existem também para outros setores de atividade, com vários grupos de indivíduos e organizações a poder beneficiar dessa disponibilidade da informação.

Por exemplo as organizações da sociedade civil, para as quais o acesso aos dados abertos pode aumentar a qualidade dos serviços que prestam e que, em alguns casos, complementam os serviços oferecidos pelo Estado.

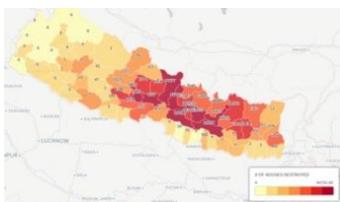
Nesse sentido, a informação detida pela Administração Pública pode ser bastante útil para estas organizações, uma vez que lhes permite conhecer melhor o seu público-alvo, a concentração geográfica dos serviços que prestam, a evolução do setor, etc. Possibilita ainda que coloquem à prova seus serviços e que possam medir os seus resultados, aumentando cada vez mais a qualidade de suas atividades.

Embora já existam inúmeros setores de atividade para os quais os dados abertos já estão a criar benefícios sociais e económicos, ainda é esperado que muitas outras aplicações se tornem possíveis graças ao crescimento desta tendência.

Mais e melhor informação “aberta” gera novas combinações de dados, que por sua vez dão origem a novas leituras e novos conhecimentos, inspirando ideias para novos campos de aplicação.

Outros exemplos de utilização dos dados abertos

AJUDA HUMANITÁRIA



Mapa interativo do terremoto no Nepal

O projeto *open data* “Code for Nepal” desenvolveu um mapa interativo que reflete os efeitos do terremoto do Nepal, distribuindo na geografia do terreno a dimensão da catástrofe em termos do número de mortos, feridos e casas danificadas. Um exemplo de como os dados podem ajudar a determinar os pontos onde a ajuda mais se impõe, tornando assim a ajuda humanitária mais efetiva.

<http://www.codefornepal.org/2015/04/map-of-reported-nepalearthquake-deaths/>



The Humanitarian Data Exchange

Saber a qualquer momento quantas camas estão disponíveis para o tratamento do ébola, quantos refugiados há num país ou qual o índice de pobreza de uma determinada cidade pode ser determinante para acelerar a resposta a uma crise humanitária. Foi precisamente para partilhar dados entre os diversos atores que, em todo o mundo, são responsáveis pela ajuda humanitária que nasceu a plataforma “Humanitarian Data eXchange” (HDX), desenvolvida e mantida pela ONU. Agregando quase 1.400 *datasets* provenientes de mais de 130 fontes distribuídas por 243 países, esta plataforma permite a partilha de dados entre os diversos atores responsáveis pela ajuda humanitária, incluindo as diversas agências da ONU, ONGs e entidades governamentais. <https://data.hdx.rwlab.org/>



disasters.data.gov

Endereçar as necessidades das equipas que respondem a situações de emergência e catástrofe é o grande propósito do portal de dados abertos “disasters.data.gov”, da responsabilidade da Casa Branca dos Estados Unidos da América. O portal integra *datasets*, ferramentas e recursos destinados à comunidade das equipas e instituições que respondem às situações de emergência. <http://www.data.gov/disasters/>

COMBATE AO CRIME ORGANIZADO



[NarcoData](http://narcodata.animalpolitico.com/)

O projeto de jornalismo de dados “NarcoData” juntou dois portais da América Latina para fazer um retrato do crime organizado no México e revelar as estruturas dos grandes grupos que têm dominado a região nas últimas quatro décadas. O conteúdo disponibilizado em “NarcoData” combina o texto com elementos de visualização e vídeo para demonstrar o crescimento do crime organizado ao longo do tempo. A iniciativa resulta da colaboração entre o portal ativista mexicano Animal Político e a plataforma chilena de jornalismo de dados Poderopedia. <http://narcodata.animalpolitico.com/>

3. COMO ABRIR OS DADOS

Falámos já sobre o que são dados abertos e porquê abrir os dados e foram dadas algumas pistas sobre como abrir os dados. Neste capítulo vamos aprofundar um pouco mais o que fazer para que os dados sejam efetivamente abertos, mostrando também qual a importância dos *standards* e as questões que levantam, pois esta é uma questão incontornável. Vai também poder ler sobre algumas das boas práticas nesta área.

Quando uma entidade pensa em abrir os seus dados pode focar-se em cinco fases essenciais: planejar devidamente o processo de abertura, saber que dados quer e deve abrir, disponibilizar efetivamente os dados de acordo com alguns princípios e regras, a estratégia de divulgação dos seus dados abertos, e por último, mas não menos importante, a manutenção dos dados.

Cinco fases para a abertura de dados



3.1. Planeamento

A opção pela disponibilização de dados abertos não pode ser encarada de ânimo leve. Deve ser sustentada por uma reflexão estratégica quanto ao tipo de informação a disponibilizar e sobre o compromisso que a entidade está disposta a assumir com a sociedade.

Além da orientação das políticas da entidade face aos dados que detém, a estratégia de disponibilização deve privilegiar o interesse e pertinência dos dados para o cidadão em geral ou para setores da sociedade específicos.

A decisão pela publicação de conjuntos de dados (*datasets*) deve ser assumida como um compromisso para com o cidadão, devendo ser sempre orientada por alguns princípios fundamentais, por forma a potenciar ao máximo a utilização dos dados e de modo a não defraudar as expectativas dos utilizadores que “investem” esforços na sua análise e no desenvolvimento de aplicações de acesso aos dados.

Nesta reflexão estratégica devem ser ponderadas preocupações distintas face a dois tipos de *datasets*:

- **Constantes**

Datasets “intemporais” sem continuidade regular (ex. Diretórios de entidades, serviços, pontos de atendimento, lojas, locais de interesse, etc.)

- **Séries**

Conjuntos de dados com espectro temporal (mensal, anual, etc.), com continuidade regular em novos *datasets* (ex. Indicadores de atividade, financeiros, etc...)

Antes de partir para a publicação de um *dataset*, ou de uma série de *datasets*, a entidade pública deve ponderar, com segurança e realismo, se tem condições de assumir o compromisso de atualização, disponibilidade e continuidade dos dados, com uma regularidade e frequência constantes.

Ou seja, depois de analisado o portfólio de informação disponível na instituição e identificados os dados a disponibilizar, será determinante criar todas as condições internas (humanas, técnicas e de processos) que permitam responder com segurança às seguintes questões:

- Os dados identificados estão em formatos abertos e reutilizáveis?
- Há condições internas para atualizar ou dar continuidade aos dados que vão ser disponibilizados?
- Com que frequência e regularidade vão ser atualizados ou ter continuidade?
- Quem vai assumir a responsabilidade pela disponibilização, atualidade e continuidade dos dados abertos em cada departamento?

Para garantir o sucesso da iniciativa de abertura de dados será importante definir claramente o modelo de governação global da iniciativa e, para cada conjunto ou série de dados, os workflows organizacionais, o plano de atualização/publicação dos dados e a responsabilidade de cada departamento.



Boa Prática: Desenvolver um plano de publicação de dados abertos

A necessidade de identificar conjuntos de dados para publicação, bem como os requisitos das partes interessadas (internas ou externas), é otimizada pelo desenvolvimento de um plano que leve em conta as necessidades dos vários agentes que estão envolvidos ou beneficiam das políticas de dados abertos.

Develop an Open Data Publication Plan. Disponível em: <https://www.w3.org/2013/share-psi/bp/odpp/>

3.2. Escolha dos dados a abrir

Tendo em conta a diversidade de entidades públicas, da informação que detêm e do público que servem, torna-se difícil definir princípios gerais para identificação de conjuntos de dados a abrir, que se apliquem a todas as entidades. No momento de selecionar os dados a disponibilizar, cada entidade deverá analisar detalhadamente o perfil e preferências do seu público e ponderar o interesse social, a utilidade setorial e o potencial económico dos dados.

As boas práticas de dados abertos indicam que um conjunto de dados é de relevância quando tem potencial para ser explorado por um grande número de utilizadores ou quando o benefício potencial da sua utilização é elevado.

Regra geral, são especialmente apreciados pelos utilizadores dados que privilegiem preocupações como:

- **Acesso** às entidades, serviços disponíveis e pontos de atendimento;
- **Transparência** das contas, atividades e recursos da entidade;
- **Utilidade pública** para o quotidiano do público-alvo;
- **Monitorização** de um tema, setor ou área de interesse público.

Nos dados a abrir deverá ser dedicada uma atenção e esforço especiais à georreferenciação da informação, porque eleva significativamente o interesse e o potencial de utilização dos dados abertos. Ao enriquecer um conjunto de dados com campos de referência geográfica, a entidade está a abrir o potencial para o desenvolvimento de aplicações que permitem visualizar os dados sobre o mapa. E estas são aplicações bastante apreciadas quer por utilizadores finais quer pela comunidade de desenvolvimento, porque permitem navegar nos dados de forma mais intuitiva e agradável e porque impulsionam a criação *apps* para dispositivos móveis.

Para a reflexão sobre os dados a abrir poderá ser determinante o envolvimento da comunidade, sobretudo os utilizadores com maior potencial de utilização dos dados.

Procure integrar desde o início do processo o ponto de vista dos utilizadores atuais e/ou dos potenciais interessados na informação que vai disponibilizar, sejam estes cidadãos, empresas, organizações não-governamentais ou *developers* (programadores). Procure saber quais as suas necessidades, expectativas e qual a utilização que preveem fazer dos dados que irá abrir, como por exemplo reutilizá-los ou cruzá-los com outras fontes de informação para gerar novas bases.

Essa consulta à comunidade pode ser feita através do pedido direto de sugestões para bases de dados, por exemplo através de um anúncio na Internet (para um público mais alargado) ou através de correio electrónico (no caso dos utilizadores habituais da informação que detém). Outra possibilidade é organizar eventos como uma *Hackathon* ou uma conferência, convidando um leque o mais abrangente possível de utilizadores para consultá-los acerca da informação que mais lhes interessa. Mais adiante neste capítulo, na área dedicada à divulgação, falaremos um pouco mais em pormenor sobre este tipo de eventos, uma vez que também são bastante utilizados para a divulgação pró-ativa da abertura de dados.

Pode também ser feita uma consulta mais direcionada, em que se disponibiliza uma pequena lista com alguns exemplos de conjuntos de dados que podem ser abertos, para tentar identificar quais são os mais importantes para o público-alvo. Há grupos dentro da sociedade civil que têm interesses particulares em determinados tipos de dados, como é, por exemplo, o caso dos jornalistas, que, com o crescimento do jornalismo de dados, trabalham também cada vez mais os dados como fontes de informação para a produção jornalística. Neste caso, podem ser realizados eventos específicos direcionados para este público-alvo.



Boa prática: disponibilizar um inventário de dados geridos

A disponibilização de um catálogo de informação referente à gestão de dados de cada organismo pode ser útil para recolher feedback sobre que datasets abrir primeiro.

Publish overview of managed data. Disponível em: <https://www.w3.org/2013/share-psi/bp/pomd/>

3.3. Tipos de utilizadores

Podem ser muito variados os utilizadores com interesse em explorar os dados abertos para os fins mais diversos, desde a utilização mais simples até à análise profissional ou a transformação dos dados em novos produtos e serviços electrónicos.

- **Administração Pública**

Como já referimos, a própria Administração Pública como parte interessada na utilização de dados abertos entre ministérios ou sectores diferentes, contribuindo para um acesso e distribuição da informação mais práticos e que funcione verdadeiramente em rede (sem necessidade de formalizar autorizações ou serviço bilaterais);

- **Empresas / empreendedores**

É no segmento das empresas que recai muito do potencial dos dados abertos, sobretudo para fins comerciais. Olhando para os dados numa perspetiva de negócio e tirando partido das tecnologias de informação e comunicação, as empresas têm condições de transformar o portfólio de dados abertos em novos produtos comerciais, seja um novo serviço *online* ou uma *app* acessível no telemóvel. Aqui há a destacar o surgimento de vários negócios e startups tecnológicas com base nos dados abertos.

- **ONGs/ativistas**

Os organismos da sociedade civil são outro segmento com enorme potencial de consumo dos dados abertos, para fins didáticos, monitorização de indicadores, etc. Projetos nacionais como o Pordata, Transparência Hackday Portugal ou Transparência e Integridade são alguns exemplos de exploração de dados abertos por Organizações Não Governamentais. A nível internacional, são já muitos os exemplos de projetos de ONGs que exploram os dados abertos para melhorar as condições de vida das populações mais desfavorecidas.

- **Jornalistas**

A nível internacional, são vários os casos de empresas de média que apostam na investigação jornalística baseada nos dados abertos. O investimento no jornalismo de dados por parte de meios de comunicação como o “The Guardian”, “La Nación” ou “ProPublica” são disso um bom exemplo. De salientar ainda o prémio anual “Data Journalism Awards” promovido pela Global Editors Network.

- **Academia**

O meio académico é parte naturalmente interessada no acesso à informação disponibilizada pelo sector público, aprofundando os seus recursos em cursos, projetos de investigação, projetos em colaboração com empresas, etc.

- **Público geral**

Os dados abertos, dadas algumas das suas características (disponíveis em *bulk* e em formatos passíveis de serem lidos por máquinas) não estão otimizados para ser consultados diretamente pelos cidadãos. Nesse sentido, e porque os dados abertos trazem enormes ganhos para toda a sociedade, é importante que quem disponibiliza dados possa assegurar também a disponibilização de mecanismos que permitam

visualizar a informação de forma inteligível, algo que a maioria dos portais de dados abertos permite.



Boa Prática: Estabelecer um ecossistema de Dados Abertos

A importância de estabelecer uma rede dinâmica que envolva as várias partes interessadas na disponibilização e reutilização de informação do sector público.

Establish an Open Data Ecosystem. Disponível em: <https://www.w3.org/2013/share-psi/bp/eode/>

3.4. Dados prioritários

Em junho de 2013, os líderes do G8 assinaram o *Open Data Charter*⁹, onde são estabelecidos os princípios estratégicos que todos estes Governos se comprometeram a seguir, na senda da abertura dos Governos.

Neste *Open Data Charter*, no capítulo 6.2, é apresentada a medida 2 que enumera 14 áreas em que os dados são considerados de grande valor e nas quais urge abrir os dados. Segundo o site do Governo do Reino Unido, a abertura de dados nestas áreas pode ajudar a abrir o potencial económico dos dados abertos, a apoiar a inovação e a fornecer uma maior responsabilização, por parte dos atores dos Governos e setor público.

É claro que estes dados não fazem sentido para todas as organizações e que possivelmente algumas das organizações e entidades ainda não estão prontas para abrir estes dados, mas conhecer quais são os dados considerados prioritários pode ser um bom ponto de partida.

Abaixo apresentamos a tabela do Open Data Charter, traduzida e um pouco adaptada, onde encontra as 14 categorias e exemplos dos tipos de dados que podem ser considerados em cada uma delas.

Categoria G8	Exemplos de <i>datasets</i>
Empresas	Registos das empresas.
Crime e justiça	Estatísticas do crime, segurança.
Monitorização Planeta Terra	Meteorologia/tempo, agricultura, florestas, pesca e caça.
Educação	Listagem de escolhas, desempenho das escolas, competências digitais.
Energia e Ambiente	Níveis de poluição, consumo de energia.
Finanças e contratação	Gastos com a contratação, contratos públicos, convites à apresentação de propostas,

⁹ <https://www.gov.uk/government/publications/open-data-charter>

Categoria G8	Exemplos de <i>datasets</i>
	orçamentos locais, orçamento de Estado (planeado e gastos efetivos), dados sobre transações comerciais internacionais.
Geospacial	Topografia, códigos postais, mapas nacionais, mapas, locais.
Desenvolvimento Global	Segurança alimentar, indústrias extrativas, terra/territórios.
Responsabilização do Governo e Democracia	Pontos de contato com o Governo, resultados das eleições (nacionais e locais), legislação e estatutos, vencimentos.
Saúde	Dados relativos às receitas médicas, dados de desempenho, locais onde são efetuadas cirurgias.
Ciência e investigação	Dados do genoma, atividade de investigação e educacional, resultados de ensaios.
Estatísticas	Dados utilizados para produzir estatísticas oficiais incluindo os Censos, inquéritos por amostragem e dados administrativos, por exemplo, dados que incluem o PIB, competências e desemprego.
Mobilidade Social e Segurança Social	Habitação, proteção na área da saúde, benefícios no desemprego.
Transportes e Infraestruturas	Horários dos transportes públicos, pontos de acesso à internet de banda larga e a sua penetração.



Boa Prática: Critérios para os *datasets*

*Esta boa prática define um conjunto de regras que podem ser usadas pelas entidades públicas para dar prioridade à publicação de determinados *datasets*. O objetivo é estabelecer um conjunto de critérios que permitam identificar “*datasets* de elevado valor”, tendo em conta a probabilidade de reutilização dos dados abertos, e ajudar as entidades governamentais a decidir sobre os *datasets* de publicação mais prioritária.*

Share PSI BP - Dataset Criteria: <https://www.w3.org/2013/share-psi/bp/dc/>

3.5. Disponibilização dos dados

O suporte de publicação é um dos aspetos essenciais do processo de abertura dos dados. Dependendo dos meios, recursos humanos e financeiros da instituição, a publicação dos dados abertos pode começar por uma solução mais simples, acrescentando uma área no atual site institucional, ou evoluir para um suporte mais complexo baseado numa plataforma de catálogo de dados abertos.

O primeiro passo pode passar pelo atual site ou portal institucional. Para a entidade que está a começar a disponibilização de dados abertos, a opção mais simples poderá passar por criar uma página ou área no site institucional, com referência aos dados disponibilizados e links para o download dos *datasets* (normalmente tabelas) nos formatos mais usados (.csv, .xls, outros).

Para as entidades que estão em condições de fazer uma aposta numa solução dedicada à gestão e publicação dos dados abertos são várias as opções de plataformas de catálogos de dados abertos.

Uma plataforma de catálogo de dados abertos consiste numa ferramenta *web* que proporciona o acesso a *datasets* sob licenças *open data*, fornecendo uma “porta de entrada” para os utilizadores que querem pesquisar, navegar e descarregar os dados disponíveis.

Dependendo da plataforma escolhida, os catálogos de dados abertos podem permitir aos utilizadores as seguintes funcionalidades:

- Pesquisar *datasets* disponíveis;
- Filtrar os *datasets* por temas;
- Fazer o *download* dos dados em formatos abertos;
- Usar APIs (*Application Programming Interface*) para programação de aplicações;
- Verificar claramente os termos de utilização (licenciamento);
- Ver a meta informação e documentação associadas ao *dataset*;
- Obter visualizações sobre os dados (gráficos, mapas, etc.);

Plataformas Open Data: Algumas opções tecnológicas

CKAN

Plataforma de catálogo de dados *open-source* (Python) suportada pela Open Knowledge Foundation (<http://okfn.org>). Esta plataforma pode ser instalada em alojamento próprio ou fornecida em modo de serviço *hosting* por uma taxa mensal.

URL: <http://ckan.org/>

DKAN

A plataforma DKAN, promovida pela Nucleo e desenvolvida a pensar na compatibilidade com a CKAN, é uma solução *open-source* baseada no sistema de gestão de conteúdos Drupal, podendo ser uma opção atrativa para as entidades que já têm sites/portais desenvolvidos em Drupal.

URL: <http://www.nucleo.com/dkan/>

Junar

Fornecido em modo de serviço em *cloud* (*cloud-based SaaS*), esta plataforma permite gerir os dados sob a infraestrutura da Junar.

URL: <http://www.junar.com>

Open Government Platform (OGPL)

À semelhança da DKAN, a plataforma OGPL é baseada no sistema de gestão de conteúdos Drupal. Esta plataforma resultou do desenvolvimento conjunto dos governos da Índia e EUA.

URL: <http://ogpl.gov.in>

Socrata

Socrata é uma solução de catálogo *open data* fornecida como serviço *online* (SaaS), fornecendo API, catálogo e ferramentas para manipulação dos dados. Uma das características distintivas da plataforma Socrata é o facto de permitir aos utilizadores criarem visualizações personalizadas sobre os dados disponibilizados.

URL: <http://www.socrata.com>



Boa Prática: Criar um portal público de dados abertos

Os portais de dados abertos podem facilitar a distribuição de dados abertos, permitindo a facilidade de acesso e pesquisa a múltiplos conjuntos de dados.

Establish Open Government Portal for data sharing. Disponível em:

<https://www.w3.org/2013/share-psi/bp/portal/>

3.6. Estratégia de divulgação

Após abrir os dados é recomendável algum investimento na divulgação da iniciativa de abertura de dados, no sentido de dar destaque ao esforço de abertura e transparência da instituição e alertando os potenciais utilizadores para a importância e potencial dos dados que passaram a ficar disponíveis em formatos abertos.

Claro que é esperado que a sociedade civil também dê resposta e participe na criação de valor à volta dos dados abertos, mas em muitas áreas, a entidade que abre os dados tem de ter um papel ativo nesta sensibilização, pelo menos numa primeira fase.

Há várias formas de divulgação da abertura de dados. Antes de mais é imperativo que a plataforma onde são disponibilizados os conjuntos de dados tenha um mecanismo de sindicância (RSS) para que qualquer cidadão possa subscrever as atualizações automáticas.

Os contactos diretos com organizações ou indivíduos e dados na sua atividade (académicos, jornalistas, etc.), pode também fazer parte da estratégia. Para tal é importante ter um bom conhecimento do público ao qual interessam os dados que estão a ser disponibilizados, criando também bases de dados com contatos que poderão ser utilizados em possíveis envios de correio eletrónico.

A divulgação de informação em *web sites* ou redes sociais relevantes, como por exemplo, comunidades de *developers* é de extrema importância. Estabelecer um bom contato com estes públicos através das redes sociais, criando grupos de discussão sobre a temática ou usando grupos ou comunidades já existentes para veicular informação sobre novos dados abertos nestas plataformas faz parte de uma boa estratégia de divulgação.

Os clássicos comunicados de imprensa também podem resultar, sobretudo se estivermos a falar de conjuntos de dados bastante importantes, ou nunca antes abertos, por exemplo. Aqui,

mais uma vez, é importante ter uma boa base de dados de contatos para os quais enviar a informação.

A organização de eventos como *Hackatons* ou conferências, já citadas neste capítulo como fazendo parte de uma possível estratégia para perceber quais os dados que mais interessam à sociedade civil que sejam abertos, são mais uma possibilidade que pode ser acrescentada na estratégia de divulgação. Neste caso, podem acabar por servir os dois propósitos. E eles podem ser organizados pelas próprias entidades que abrem ou querem abrir os seus dados, ou então as entidades podem criar parcerias com alguns destes eventos que já existem.

A divulgação é de extrema importância para um projeto de dados abertos e deve ser tida sempre em consideração na sua planificação.

O papel das Hackathons

O que é uma *Hackathon*?

Uma *Hackathon* é no seu sentido mais literal uma maratona de programação. Normalmente este tipo de eventos reúne programadores, designers e outros profissionais ligados ao desenvolvimento de software para uma maratona de programação, cujo objetivo é desenvolver um *software* que atenda a um fim específico ou projetos livres que sejam inovadores.

No caso das *Hackathons* na área dos dados abertos, o objetivo é utilizar os dados para desenvolver produtos/serviços úteis e interessantes. Elas servem também para promover e divulgar, por exemplo, o trabalho feito com a abertura de dados e os *datasets* disponíveis, e como já dissemos criar *apps* e serviços e podem servir para validação e teste de dados.

Em Portugal temos alguns exemplos interessantes de *Hackatons* bem-sucedidas, sendo muitas organizadas por estabelecimentos de ensino, organizações sem fins lucrativos e empresas da área das Tecnologias de Informação e Comunicação. Há também algumas organizadas por entidades da administração pública, como é caso da Câmara Municipal de Lisboa.

Um caso de Lisboa

A “TM Fórum Action Week” decorreu em Lisboa no início de fevereiro de 2016 e desafiou os participantes a conceberem serviços tecnológicos inovadores e criadores de valor para as cidades. Com acesso à plataforma “Open Eco System”, onde são disponibilizados pelos parceiros da TM Fórum vários dados abertos da cidade, os participantes foram desafiados a conceber novos e inovadores serviços tecnológicos, criadores de valor para as cidades.

A Câmara Municipal de Lisboa aproveitou o evento para apresentar a sua Plataforma Geodados. Ainda em fase de testes, esta plataforma disponibiliza dados do município que podem ser usados livremente. Nesta plataforma estão disponíveis mais de 190 itens distribuídos por 20 áreas, incluindo transportes, meteorologia, escolas, restaurantes e parques.

Ligações úteis:

Hackathon Guide: <https://hackathon.guide/>

The Hack Day Manifesto: <http://hackdaymanifesto.com>

3.7. Manutenção dos dados

Uma iniciativa de abertura de dados está longe de terminar com o lançamento público e a disponibilização dos dados abertos ao público. O sucesso do projeto de dados abertos muito depende da estratégia, atitudes e práticas que a entidade assumir ao longo do tempo. E são vários os desafios e exigências colocados ao promotor de uma iniciativa de dados abertos, a começar pela disponibilidade, atualização, continuidade e qualidade dos dados

À semelhança das restantes áreas da informação pública, também os dados abertos devem ser orientados pelo princípio da **disponibilidade contínua**. Ou seja, uma vez disponibilizado um conjunto de dados, a entidade pública deverá garantir que esses dados estão sempre disponíveis na hiperligação aos ficheiros originais.

A **atualidade** dos dados é um dos fatores com maior impacto na relevância e interesse do cidadão face aos *datasets* já disponibilizados pela instituição. Ao disponibilizar um conjunto de dados, uma entidade deve assumir o compromisso de vigiar e manter os *datasets* já carregados em constante atualização. Esta é uma preocupação que assume especial importância nos *datasets* constantes que são relativamente estáticos, como listagens de entidades, pontos de atendimento, lojas, etc.

Ao iniciar a disponibilização de uma série de dados sobre um determinado tema, a entidade pública está a abrir ao cidadão uma porta de informação que se espera fiável e duradoura, sob pena de deitar por terra todo o interesse do utilizador e os esforços que ele investiu na análise e desenvolvimento de aplicações.

Sobretudo nas séries, com espectro temporal nos dados (mensal, trimestral, anual, etc.), é essencial garantir a continuidade dos dados, acrescentando novos *datasets* ao tema. Tudo para que não seja interrompido um fluxo de informação que alimentou no utilizador expectativas de continuar “já a seguir”.

A par da **continuidade**, a regularidade dos dados é outro fator determinante para fidelizar o interesse e entusiasmo do utilizador face a uma série de *datasets*. Se a entidade começar a disponibilizar uma série de dados mensais, convém que o utilizador não tenha de esperar mais de um mês para ter os dados atualizados. Na prática, trata-se manter o fluxo de dados

com uma regularidade previsível, seja na frequência com que os *datasets* são disponibilizados seja no espectro temporal dos dados de cada *dataset*.

Outro aspeto importante é a **qualidade** dos dados, de grande impacto na leitura, manipulação e esforço de desenvolvimento que o utilizador “investe” nos *datasets* disponibilizados pela entidade pública. É de todo o interesse fornecer ao público dos dados uma “matéria-prima” compreensível, intuitiva e em perfeitas condições para ser trabalhada e manipulada. Esta é uma preocupação de extrema importância que, num *dataset*, abrange desde o título e a meta informação até ao detalhe dos nomes dos campos e diferentes tipos de dados.



Share-PSI 2.0

3oa Prática: Estabelecer canais de feedback para melhoria da qualidade

A disponibilização de canais de feedback que possibilitem o reporte de erros, inconsistências, incoerências ou informação incompleta, por parte da comunidade de reutilizadores, pode contribuir para a melhoria significativa da informação disponibilizada.

Enable feedback channels for improving the quality of existing government data. Disponível em:

<https://www.w3.org/2013/share-psi/bp/ef/>

3.8. Exemplos de iniciativas de dados públicos abertos

Em Portugal

Dados.Gov

O Dados.Gov é o portal de dados abertos português, publica e agrega a informação produzida pela Administração Pública em formatos abertos que possam ser lidos e reutilizados por qualquer cidadão. O seu principal objetivo é facilitar o acesso à informação pública, promover a transparência da Administração Pública e ajudar na criação de serviços públicos eletrónicos pela sociedade civil.

Ao promover a reutilização da informação pública, o Dados.Gov pretende também fomentar a existência de novos serviços destinados aos cidadãos em geral.

Produto da Agenda Digital, lançada pelo Resolução do Conselho de Ministros n.º 91/2010, quase no final de 2010, onde é mencionada a criação do Dados.Gov, é em novembro de 2011 que é lançada a sua versão beta, então o portal tinha 100 *datasets*.

O portal é um projeto da AMA - Agência para a Modernização Administrativa, IP, que conta com a participação de inúmeros organismos da Administração Pública Central, Regional e Local como parceiros.

<http://www.dados.gov.pt>

Portal de Transparência do Serviço Nacional de Saúde

O Portal da Transparência é uma iniciativa de dados abertos do Ministério da Saúde que tem por objetivo disponibilizar os conjuntos de dados relativos às operações e transações das diversas entidades que integram o Sistema Nacional de Saúde (SNS).

São mais de 50 os *datasets* disponibilizados no catálogo de dados logo no arranque do Portal da Transparência, agregando dados de diferentes entidades do SNS em vários domínios. A navegação por este portfólio de dados abertos da saúde é facilitada graças às ferramentas de ordenação e filtragem, que permitem consultar os *datasets* em função do tema, palavra-chave, entidade, tipo de vista ou data de alteração.

O Portal está inserido no portal do Serviço Nacional de Saúde.

<https://transparencia.sns.gov.pt>

Dados abertos Biblioteca Nacional de Portugal - opendata bnportugal

A Biblioteca Nacional de Portugal (BNP) fornece no site “opendata bnportugal” informação e acesso aos conjuntos de dados que disponibiliza livre e gratuitamente. Nestas páginas *web* encontra todo o catálogo bibliográfico da BNP, da Base Nacional de Dados Bibliográficos - PORBASE, a Bibliografia Nacional Portuguesa (desde 1931) e Biblioteca Nacional Digital, entre outros.

<http://opendata.bnportugal.pt>

Lisboa Aberta – Dados abertos do Município de Lisboa

A Câmara Municipal de Lisboa lançou, em fevereiro de 2016, um *upgrade* do seu portal de dados abertos. A nova plataforma “Lisboa Aberta” apresenta-se como um portal de dados abertos que disponibiliza informação em vários domínios da vida da capital portuguesa, incluindo transportes, serviços, cultura ou dados estatísticos.

O novo portal conta com a participação de vários atores da cidade, que se juntaram em parceria ao projeto, entre eles contam-se a Fertagus, Administração do Porto de Lisboa, ANA, MEO, EDP, Associação de Turismo de Lisboa, Instituto Nacional de Estatística, EPAL, Transportes de Lisboa ou EMEL.

Estas parcerias de abertura de dados da cidade são orientadas por três grandes objetivos: aumentar a participação e o envolvimento dos cidadãos, promover a capacitação e a inclusão digital, estimular a partilha de dados e a inovação aberta.

<http://dados.cm-lisboa.pt/>

Na Europa

Portal Europeu de Dados Abertos

A Comissão Europeia lançou, no final de 2015, a primeira versão do Portal de Dados Europeu. O novo portal agrega metadados de informação do setor público disponível em mais de 28 portais de dados abertos dos países europeus, incluindo o portal nacional de dados abertos Dados.Gov. No total, estão já referenciados no Portal de Dados Europeu mais de 240 mil *datasets* dispersos por 34 países.

<http://www.europeandataportal.eu/>

European Union Open Data Portal

Desenvolvido pela Comissão Europeia, este portal permite aceder aos dados produzidos pelos próprios institutos e organismos que compõem a União Europeia (UE), que são assim disponibilizados para poderem ser utilizados sem qualquer tipo de restrição por qualquer pessoa ou entidade.

3.9. Dados.Gov: O portal nacional de dados abertos

A par da disponibilização dos dados abertos nos suportes próprios da instituição, ou em alternativa a esta opção, a publicação no portal nacional de dados abertos, Dados.Gov, é uma condição importante para impulsionar a distribuição e disseminação dos dados abertos a um público mais alargado, de entidades públicas, cidadãos e empresas.

Inserido nas responsabilidades da Agência para a Modernização Administrativa, IP (AMA) na promoção do Governo Aberto, o Dados.Gov é o ponto central na estratégia nacional de dados abertos, constituindo-se como a porta central de acesso *online* ao repositório dos dados abertos disponibilizados pelas várias instituições da Administração Pública em Portugal. O

Dados.Gov disponibiliza vários *datasets*, provenientes de organismos diferentes, que qualquer cidadão, empresa ou entidade, que pode descarregar, modificar e reutilizar para diversos fins.

Hoje em dia, a plataforma Dados.Gov partilha centenas de conjuntos de dados abertos de muitas dezenas de instituições públicas e garante o suporte de dados a várias aplicações de grande relevância nacional, como o Portal de Transparência Municipal ou o Mapa do Cidadão.

A nível internacional, a AMA é a entidade nacional que representa o Governo português nos vários fóruns de alto nível, grupos técnicos e iniciativas dedicadas à temática *open data*.

Fruto desta longa experiência nos vários domínios da abertura de dados, a equipa da AMA tem todas as condições e a disponibilidade para prestar às entidades públicas o apoio e a consultoria especializada nas várias fases do projeto de abertura de dados.

Por tudo isto, a AMA tem todas as condições para dar uma ajuda importante às instituições públicas que, estando sensibilizadas para o imperativo dos dados abertos, ainda têm dúvidas quanto à melhor estratégia e opções para a abertura dos seus dados, em linha com as boas práticas e recomendações internacionais.

No desenvolvimento de uma iniciativa de dados abertos, a AMA pode ajudar a entidade pública a encontrar respostas para dúvidas como:

- Que dados disponibilizar?
- Como abrir os dados?
- Que estratégia de implementação do projeto?
- Como conquistar o público?

Contatos

AMA

URL: www.ama.pt | Email: ama@ama.pt | Tel.: 21 723 12 00

Dados.Gov

URL: www.dados.gov.pt

4. REUTILIZAR OS DADOS

A abertura de dados está a acontecer um pouco por todo o mundo e está a transformar a forma como os dados são recolhidos, publicados e utilizados. No entanto, quando uma entidade inicia o caminho de disponibilização de dados abertos, as decisões não se resumem apenas à escolha dos dados a abrir ao público.

Selecionada a informação a disponibilizar, definidas competências e os *workflows* organizacionais indispensáveis à abertura de dados, há que refletir e estruturar bem os detalhes mais técnicos da disponibilização dos dados.

Que formatos de ficheiros usar? Que normas seguir na estruturação dos dados? Que especificações adotar para a metainformação associada aos dados? E qual é o licenciamento mais adequado para assegurar que os dados podem ser realmente reutilizados sem comprometer os interesses da instituição fornecedora?

Estas e outras questões de detalhe associadas à disponibilização de dados em formatos abertos são indispensáveis ao sucesso de qualquer iniciativa *open data*. Falamos aqui de *standards*, a necessidade de definir e documentar modelos normalizados para a recolha, publicação e utilização dos dados abertos.

Um imperativo de normalização nos dados abertos que Steve Adler, Chief Information Strategist da IBM e co-Chair do *Data on the Web Best Practices Working Group* do W3C, sublinha num texto publicado no LinkedIn¹⁰.

Segundo Steve Adler, a falta de normalização dos procedimentos leva frequentemente à publicação de dados com metadados inconsistentes, com falta de documentação oficial dos processos de aprovação a apoiar a publicação e, muitas vezes, termos de utilização que resultam em conflito. Adler fala da necessidade de os dados de abertos provenientes de mais de uma fonte serem agregados, normalizados, limpos, de qualidade, autênticos e com as devidas condições de utilização.

A procura da criação de *standards* nasce da necessidade de consenso, equidade e qualidade dos dados disponibilizados.

Quando se consegue chegar a *standards* comumente utilizados obtêm-se ganhos em três áreas fundamentais: na interpretação dos dados pelas máquinas, na forma como é feito o trabalho para disponibilização dos dados, e até mesmo um pouco a jusante, na forma como

¹⁰ <https://www.linkedin.com/pulse/open-data-standards-steven-adler>

são preparados os dados para a disponibilização, e na interpretação e acesso aos dados por parte dos públicos aos quais se destinam.

Daí que seja tão importante a reflexão e decisão prévia de questões detalhe como os formatos de dados, o modelo de metadados ou a opção do licenciamento.

Escolhas e definições que devem ficar bem documentadas para que possam ser respeitadas por todos os envolvidos em qualquer uma das fases do processo de abertura de dados, desde os departamentos dispersos que têm de enviar dados até aos responsáveis pela publicação dos *datasets*.

4.1. Formatos abertos

Para quem se prepara para disponibilizar dados em formatos abertos pela primeira vez, uma das questões mais frequentes prende-se com os formatos de ficheiro. Que formatos de ficheiro usar para a disponibilização de dados abertos? Qual o ficheiro mais adequado ao tipo de dados que temos para disponibilizar? E qual o melhor para o público que vai usar os dados disponíveis?

A resposta a estas dúvidas sobre formatos deve ter como ponto de partida dois princípios essenciais dos dados abertos: os dados devem ser disponibilizados em formato aberto e legível por máquina. Ou seja:

- **Formatos abertos:** são formatos que não são proprietários e são independentes de quaisquer plataformas, podendo ser acedidos por qualquer pessoa e sem necessidade de recurso a qualquer *software* licenciado.
- **Formatos legíveis por máquinas:** são formatos que permitem que um sistema informático leia os dados contidos no ficheiro. Estes formatos são estruturados e facilmente pesquisáveis através de código.

Claro que, havendo muitas opções de ficheiros, os formatos mais apropriados dependerão sempre do tipo de dados a disponibilizar. Haverá opções mais indicadas para tabelas simples, para dados mais complexos, para dados geográficos, etc. Por exemplo, uma publicação PDF pode ter um aspeto mais agradável mas pode impor grandes limitações ao utilizador na hora de trabalhar e reutilizar a informação.

Os formatos abertos e legíveis por máquinas permitem que os dados sejam usados e editados facilmente, além de facilitar a interoperabilidade entre diferentes *datasets*.

Nesta decisão dos formatos será importante ter em conta o modelo de dados abertos (5 Estrelas da Publicação de Dados Abertos) definido por Tim Berners-Lee. As boas práticas da abertura de dados dizem que o formato escolhido deveria satisfazer no mínimo as exigências das 3 estrelas de Berners-Lee.

5 Estrelas da Publicação de Dados Abertos – Tim Berners-Lee

★	Disponível na Web, independentemente do formato, mas com uma licença aberta
★ ★	Disponível em formato estruturado e que possa ser lido por máquinas
★ ★ ★	Disponível em formato não proprietário
★ ★ ★ ★	Todos os anteriores e utilização dos standards World Wide Web Consortium (W3C)
★ ★ ★ ★ ★	Todos os anteriores e permitir a ligação a outros dados para oferecer contexto (linked data)

Exemplos de formatos abertos¹¹

- **Comma Separated Values (CSV)**

O formato CSV é um dos formatos mais populares para *datasets* por ser um dos suportes mais básicos, legível e importável por qualquer sistema ou *software*. Trata-se de um formato txt composto por: uma linha de cabeçalho com o rótulo para cada campo; e linhas de dados com os valores separados por vírgulas.

- **Tab-Separated Values (TSV)**

Semelhante ao CSV, é um formato baseado em txt mas com os valores dos dados separados por tabs. Também é um formato de dados universal e legível por máquinas.

¹¹ Fonte: Government of South Australia Open Data Process Guide https://toolkit.data.gov.au/index.php?title=Publishing_your_data#Intro_to_metadata

- **JavaScript Object Notation (JSON)**

O formato JSON usa texto de leitura humana para transmitir objetos de dados compostos por pares de valores. É usado sobretudo para transmitir dados entre um servidor e uma aplicação *Web*, em alternativa ao XML. O tamanho do ficheiro JSON é mais compacto e reduzido do que o formato XML.

- **Extensible Markup Language (XML)**

O formato XML é uma linguagem *markup* amplamente conhecida que define um conjunto de regras para a codificação de documentos num formato que é legível tanto por máquinas como por humanos. Os utilizadores criam e definem as suas próprias *tags*.

- **Rich Site Summary (RSS)**

RSS (originalmente *RDF Site Summary*) utiliza uma família de *standards* de *web feeds* para a publicação de informação atualizada com frequência: artigos de blogs, notícias, áudio, vídeo. O documento RSS, também denominado “feed”, inclui texto de resumo, ligação para a fonte e metadados, como data de publicação e nome do autor, etc.

- **ATOM**

O formato de sindicância Atom consiste numa linguagem XML usada e *feeds web*, desenvolvida como alternativa ao XML.

- **Open Document Format for Office Applications (ODF)**

O formato ODF, também conhecido como *OpenDocument*, é um formato de ficheiro baseado em XML para documentos como folhas de cálculo, gráficos, apresentações e texto. Foi desenvolvido com o objetivo de proporcionar uma especificação de formato de ficheiro aberto para aplicações office.

- **HyperText Markup Language (HTML)**

Formato de ficheiros usado para a formação de informação *web*.

- **Keyhole Markup Language (KML)**

O formato KML é uma linguagem XML dedicada a visualização geográfica, incluindo anotação e mapas e imagens.

- **Geography Markup Language (GML)**

O formato GML consiste numa gramática XML definida pelo *Open Geospatial Consortium* (OGC) para expressar características geográficas, servindo como linguagem de modelação para sistemas geográficos assim como formato de interligação aberto para transações geográficas na Internet.

- **GeoJson**

Formato de *standard* aberto para a codificação de bibliotecas de características geográficas simples. Inclui atributos não-espaciais usando *JavaScript Object Notation*.

Ligações Úteis

Government Service Design Manual - Choosing appropriate formats

<https://www.gov.uk/service-manual/user-centred-design/choosing-appropriate-formats.html>

Open Format Definition

<http://opendefinition.org/ofd/>

Sunlight Foundation Open Data Guidelines

<http://sunlightfoundation.com/opendataguidelines/>

W3C Best Practice - Make the data available in a language people want

[http://www.w3.org/2013/share-](http://www.w3.org/2013/share-psi/wiki/Best_Practices/Make_the_data_available_in_the_language_people_want_it)

[psi/wiki/Best_Practices/Make_the_data_available_in_the_language_people_want_it](http://www.w3.org/2013/share-psi/wiki/Best_Practices/Make_the_data_available_in_the_language_people_want_it)

4.2. Estrutura de dados

A qualidade dos dados é um dos aspetos essenciais a ter em conta na disponibilização de dados em formatos abertos. Sobretudo se tivermos em conta que os dados são abertos precisamente para serem utilizados, manipulados e explorados por utilizadores e sistemas externos.

A partir do momento em que a entidade decide passar a disponibilizar ao público um fluxo de dados num determinado tema (despesas, receitas, serviços, pontos de atendimento, etc.) tem de ter a noção que está a abrir a porta a utilizadores, programadores e empresas, que passam a investir tempo e esforço na leitura, manipulação e desenvolvimento nos *datasets* disponibilizados pela entidade.

Daí que seja de todo o interesse fornecer ao público dos dados uma “matéria-prima” compreensível, intuitiva e em perfeitas condições para ser trabalhada por pessoas ou por sistemas.

Esta é uma preocupação de extrema importância que, num *dataset*, abrange desde o título e a meta informação até ao detalhe dos nomes dos campos e diferentes tipos de dados.

- **Título**

O título será, provavelmente, um dos detalhes mais importantes de um *dataset* por ser a porta de entrada para os dados, devendo primar por ser simples, conciso, de leitura imediata e distintivo face aos restantes *datasets*. Nas listagens do portfólio de datasets, o título deverá ser capaz de orientar o utilizador, de forma clara e sem dúvidas, para os dados que mais lhe convêm. Em séries de *datasets*, contínuas e sequenciais, deverá haver uma preocupação de normalização que permita reconhecer facilmente a sequência.

- **Nomes dos Campos**

Os termos que dão nome aos campos devem ser outra preocupação a ter em conta na estruturação dos dados a disponibilizar em *dataset*. Os nomes dos campos devem ser simples mas compreensivos, não deixando dúvidas quanto aos dados que contêm.

Siglas e abreviaturas devem ser usadas com o máximo cuidado, de forma a não prejudicar a leitura do utilizador. Sempre que as siglas forem inevitáveis, recomenda-se a descrição na meta-informação. Em dados do mesmo tipo, convém usar termos e formatos idênticos (maiúsculas/minúsculas; espaço/*underscore*, etc.). Em séries de *datasets* contínuas e sequenciais, os nomes dos campos devem seguir um padrão normalizado, tanto nos termos como nos formatos.

- **Dados**

Nos dados propriamente ditos, a estruturação dos dados pelos diferentes campos deve seguir as boas práticas da estruturação de dados, incluindo o respeito pelas características próprias dos diferentes tipos de dados (texto, número, data, etc.), a desagregação de dados distintos por diferentes campos e a normalização de dados ao longo dos campos.

- **Georeferenciação**

Os dados com georeferenciação são dos mais apetecíveis à criação de aplicações que exploram o potencial dos dados abertos, permitindo traduzir os dados em pontos sobre o mapa ou rotas em dispositivos móveis.

Sempre que os *datasets* tenham uma referência local (conselho, cidade, freguesia, morada) aconselha-se acrescentar campos dedicados à informação geográfica ou criar um *dataset* relacionado com os dados de georreferenciação.

 Share-PSI 2.0	Boa Prática: Normas para dados geoespaciais
<p><i>A informação geográfica produzida e gerida no âmbito do sector público é utilizada por vários sectores de grande importância económico-social (serviços públicos, construção, sector jurídico, serviços sociais, atividades financeiras, outros). Nesse sentido, é essencial que a informação geoespacial seja partilhada de forma a poder ser reutilizada pelo maior número de organizações possível, obedecendo a uma série de normas internacionais.</i></p> <p><i>Standards for Geospatial Data. Disponível em: https://www.w3.org/2013/share-psi/bp/sgd/</i></p>	

4.3. Metadados

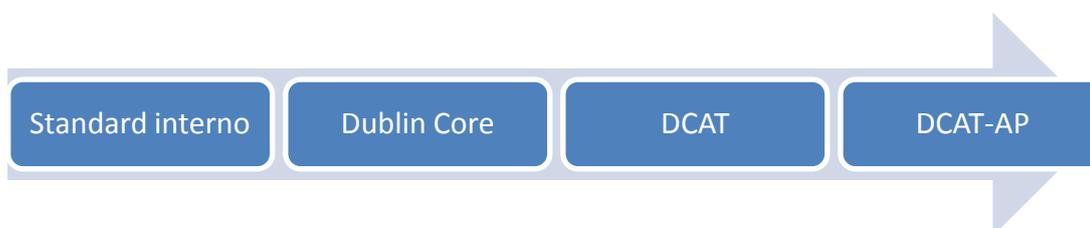
Os metadados, também referidos como metainformação ou metadata, podem parecer um pormenor secundário, mas são um detalhe de extrema importância para a utilização eficiente dos dados disponibilizados em formatos abertos.

Os metadados acrescentam ao *dataset* toda a informação descritiva sobre os dados, descrevendo o conteúdo (título, descrição, etiquetas), o formato, as limitações, a frequência e a atualização, entre outros detalhes que podem ser associados ao *dataset*. Esta informação é essencial para dar ao utilizador, ou ao sistema informático, um contexto sobre os dados que, estando bem estruturado, torna-se decisivo para a interoperabilidade com outros *datasets*.

Também aqui, os procedimentos de recolha e publicação não podem ser deixados à mercê do acaso ou do imprevisto de cada um dos intervenientes que, dentro da instituição, participa no processo de abertura de dados. Nesse sentido, convém que todo o processo seja sustentado por normas de metadados bem documentadas que sejam seguidas de forma consistente por toda a organização, independentemente da pessoa, do departamento ou da data de publicação.

Se, numa série de *datasets* sobre o mesmo tema (ex.: despesas), forem usadas descrições diferentes (despesas, custos, etc.) em cada dataset, não será difícil antecipar problemas para os utilizadores ou sistemas que já têm parametrizado a pesquisa e exploração desses dados.

As opções de normalização, ou *standards*, para os metadados podem variar muito, dependendo da complexidade e maturidade que a entidade está disposta a investir no modelo de descrição dos dados. A experiência descrita por vários projetos de dados abertos a nível internacional permite identificar vários estados na evolução do modelo de metadados:



- ***Standard Interno***

Este é geralmente o primeiro passo na evolução da metainformação, em que a entidade adota a sua própria lógica e modelo descritivo dos dados, incluindo campos, descrições, etc.

- ***Dublin Core***

O esquema de metadados Dublin Core, mais avançado, permite descrever objetos digitais, tais como vídeos, sons, imagens, textos e sites *web*, com base em XML e RDF (*Resource Description Framework*). O esquema foi desenvolvido pela *Dublin Core Metadata Initiative* (DCMI), uma organização internacional que se dedica a promover a adoção de padrões de interoperabilidade de metadados e a desenvolver vocabulários especializados para descrever fontes e recursos *web*, por forma a facilitar e agilizar os processos de pesquisa e recuperação de informação.

- ***DCAT***

No caminho de evolução dos metadados, o ideal seria que todas as entidades públicas evoluíssem para o *standard Data Catalog Vocabulary* (DCAT), uma norma de metadados mais completa e de alto nível que é usada para descrever todos os dados públicos na Europa. Pretende-se que a sua utilização contribua para tornar os dados públicos pesquisáveis pelos diferentes países e pelos diversos setores.

- ***DCAT-AP***

A nível europeu também está a ser desenvolvida a especificação “DCAT Application profile for data portals in Europe” (DCAT-AP) para a descrição de *datasets* setoriais na Europa. A sua utilização básica pretende permitir a pesquisa de *datasets* por vários portais e tornar os dados do setor público mais pesquisáveis por todos os países e

setores, através do intercâmbio de descrições de *datasets* entre os portais de dados abertos dos vários países europeus. Em fevereiro de 2015, o programa ISA da Comissão Europeia iniciou o processo de revisão da especificação DCAT-AP, com base na experiência adquirida desde o desenvolvimento da norma em 2013.

Independentemente da opção escolhida para os metadados, e do seu estado de evolução e maturidade, será sempre importante definir claramente e documentar que metainformação deverá ser recolhida para os diferentes *datasets*. E, desde o início da abertura de dados, começar a registar a metainformação de uma forma consistente em todos os dados, segundo as normas definidas.

Nesse sentido, é aconselhável que a organização vá construindo o seu próprio catálogo de metadados, integrando-o nos processos de recolha e publicação dos dados e assegurando que a metainformação vai sendo associada aos dados de forma regular e consistente. A utilização de *standards* intermédios, como o Dublin Core, pode revelar-se importante para ajudar a entidade, e o seu projeto de dados abertos, a evoluir para uma norma de alto nível, como o DCAT.

Com efeito, o *standard* DCAT será a grande meta na evolução da metainformação dos projetos de dados abertos na Europa. Ao acrescentar aos dados públicos um modelo de metainformação transversal e pesquisável por todos os países e setores na Europa, a utilização do *standard* DCAT dá um enorme potencial à pesquisa e exploração de dados a nível europeu.

Ligações úteis

DCMI Metadata Basics

<http://dublincore.org/documents/dces/>

Dublin Core Elements

<http://dublincore.org/metadata-basics/>

NISO Understanding Metadata

<http://www.niso.org/publications/press/UnderstandingMetadata.pdf>

The ODI - Marking up your dataset with DCAT

<https://theodi.org/guides/marking-up-your-dataset-with-dcat>

W3C Best Practice – Metadata

<http://www.w3.org/TR/dwbp/#metadata>

DCAT application profile for data portals in Europe

https://joinup.ec.europa.eu/asset/dcat_application_profile/description

4.4. Licenças abertas

A licença é uma condição essencial à abertura de dados. E estando na abertura a essência de um projeto de dados abertos também a licença deve ter um perfil aberto. Mas o que significa exatamente uma licença? E porque é tão importante que um projeto de dados esteja abrangido por uma licença aberta?

Quando alguém se cria algo, um trabalho original ou uma fotografia por exemplo, o seu autor obtém automaticamente direitos sobre o trabalho, podendo determinar em que moldes o seu trabalho pode ser usado por outros.

Ao aplicar a um trabalho uma determinada licença, o autor está a expressar explicitamente que direitos quer manter, ou abdicar, sobre a propriedade do seu trabalho e as condições em que outros utilizadores podem usar esse trabalho.

Para um autor que pretende disponibilizar o seu trabalho em moldes abertos, será sempre necessário aplicar uma licença aberta. Ou seja, uma licença que expresse claramente que o autor abdica de direitos de propriedade originais, para dar a outros utilizadores a possibilidade de reutilizar, modificar e partilhar o seu trabalho.

Dependendo dos direitos de propriedade a que o autor está disposto a abdicar, são várias as opções de licenças disponíveis. Para um projeto de dados abertos, a escolha da licença deverá ter em conta o potencial de exploração dos dados e o interesse do público-alvo.

Entre as licenças abertas mais populares destacam-se:

Nível da licença	Licença Creative Commons	Licença Open Data Commons
Public domain (todos os direitos abdicados)	CC0	PDDL
Atribuição	CC-by	ODC-by
Atribuição – Partilha	CC-by-sa	ODbL



Public Domain

Licença universal do titular de direito de autor e/ou de direitos conexos e/ou de direitos sobre bases de dados que pretende renunciar a todos os seus direitos sobre o trabalho, em todo o mundo.



Atribuição (by)

Nos termos desta licença a utilização da obra é livre, podendo os utilizadores fazer dela uso comercial ou criar obras derivadas a partir da obra original. Essencial é, apenas, que seja dado o devido crédito ao seu autor.



Atribuição (by-nc)

De acordo com esta licença o autor permite uma utilização ampla da sua obra, limitada, contudo, pela impossibilidade de se obter através dessa utilização uma vantagem comercial. É também essencial que seja dado o devido crédito ao autor da obra original.



Atribuição – Partilha nos Termos da Mesma Licença (by-sa)

Quando um autor opte pela concessão de tal licença pretenderá, não só que lhe seja dado crédito pela criação da sua obra, como também que as obras derivadas desta sejam licenciadas nos mesmos termos em que o foi a sua própria obra. Esta licença é muitas vezes comparada com as licenças de software livre.

4.5. Escolha da licença

Para um projeto de dados abertos, o licenciamento dos dados é essencial para conferir aos utilizadores potenciais a clareza e certeza na utilização e exploração da informação disponibilizada. E deverá respeitar um dos princípios chave da abertura de dados: garantir uma utilização sem quaisquer restrições.

Havendo várias opções, a escolha da licença deverá estar em linha com a estratégia de abertura de dados da organização e orientada pelas metas a alcançar, em termos dos objetivos de utilização dos dados e do perfil dos utilizadores a conquistar.

A licença que dá mais garantias de sucesso à utilização dos dados, para os mais variados fins, é a que menos direitos reserva, a licença *Public Domain*. Ao abdicar de todos os direitos de propriedade, esta é a licença que mais garantias dá aos utilizadores que querem explorar os dados para os transformar em aplicações e serviços, sobretudo para fins comerciais.

Exigir uma licença do tipo “atribuição”, como a que é utilizada no Portal Nacional de Dados Abertos – Dados.Gov, e que pode não ser tão atrativa para os utilizadores comerciais, porque obriga a fazer hiperligação (*link*) para o conteúdo original. No entanto, pode ser uma boa opção para ajudar a promover a iniciativa *open data*.

Já a licença do tipo “atribuição-partilha” exige aos utilizadores que publiquem os trabalhos derivados dos dados também de forma aberta. Uma opção de licenciamento que também pode afastar os utilizadores comerciais e pessoas que querem fazer lucro com a utilização dos dados.

As licenças que mantêm algumas restrições, como as “atribuição” e “atribuição-partilha”, ao afastarem os utilizadores comerciais podem diminuir o potencial de inovação e utilização.

4.6. Publicar a licença

Selecionada a opção de licenciamento mais adequada à estratégia e objetivos da iniciativa de dados abertos, a licença deve ser aplicada de forma clara se sem deixar dúvidas ao utilizador que quer explorar e reutilizar os dados.

Nesse sentido será importante adotar alguns procedimentos:

- **Identificação visual**

O tipo de licença deve estar perfeitamente visível em todo portal ou página onde são disponibilizados os dados abertos, através de logós identificativos da licença ou com referência expressa. De preferência, a identificação da licença deve ter hiperligação para conteúdo descritivo dos termos da licença.

- **Página de termos e condições**

Os termos e condições da licença aplicada aos dados devem estar descritos, de forma clara e concisa, numa área dedicada. Se a disponibilização dos dados abertos estiver integrada no *site*/portal da instituição, a licença dos dados deve ser bem destacada de forma a não se confundir com termos e condições gerais do *site*/portal.

- **Clareza e simplicidade**

A descrição dos termos e condições deve ser o mais claro e simples possível, de forma a não confundir o utilizador.

Ligações úteis

Licenças Creative Commons

<http://creativecommons.pt/cms/view/id/28/>

Licenças Creative Commons - Ferramentas

<http://creativecommons.pt/cms/view/id/14/>

Creative Commons license chooser

<http://creativecommons.org/choose/>

Open Data Commons licenses

<http://opendatacommons.org/licenses/>

Open Definition - Open Knowledge Foundation

<http://opendefinition.org/licenses/>