



LICENCIATURA

EM

CIÊNCIA DA INFORMAÇÃO

Adopção da plataforma OpenDataSoft para publicação de dados abertos

Carolina Miranda da Silva Durães

L

2024



Adopção da plataforma OpenDataSoft para publicação de dados abertos

Carolina Miranda da Silva Durães

Adopção da plataforma OpenDataSoft para publicação de dados abertos

Relatório de Projeto realizado no âmbito da Licenciatura em Ciência da
Informação, orientado pela Prof.^a Mariana Curado Malta e supervisionado pelo o

Dr. Pedro Pimenta

Faculdade de Letras e Faculdade de Engenharia
Universidade do Porto

Junho de 2024

Sumário

Sumário	4
Resumo	5
Índice de Figuras e Tabelas	6
Lista de Siglas e Abreviaturas	7
1. Introdução	9
1.1 Contexto	9
1.2 Problema	10
1.3 Objetivos do estágio	10
1.4 Organização do relatório	10
2. Metodologia e plano de trabalho	11
3. Contextualização em Ciência da Informação	12
4. Estado da Arte	14
4.1 Dados/ Dados Abertos	14
4.2 Metadados	15
4.3 Data Lake	16
4.4 Catálogo de Dados	17
4.5 ISO 37120	17
4.6 OpenDataSoft	18
4.7 DCAT	19
4.8 OpenDataSoft e DCAT	23
5. Concretização do projeto	24
4.1 Desenho da solução, apresentação, avaliação e discussão dos resultados	24
6. Conclusões e reflexão crítica	31
6.1. Lições aprendidas	31
6.2. Trabalho futuro	31
Referências bibliográficas	32
Anexos	35
Anexo 1	35
Anexo 2	37
Anexo 3	41
Anexo 4	45
Anexo 5	48
Anexo 6	50

Resumo

Este documento apresenta o relatório de estágio realizado na Câmara Municipal da Maia¹ no contexto da disciplina de projeto da licenciatura em Ciência da Informação da Faculdade de Letras e Engenharia da Universidade do Porto, apresento este relatório de estágio que descreve o trabalho realizado. O projeto teve como objetivos o desenvolvimento nas áreas de gestão de dados e dados abertos, incluindo a adaptação da plataforma OpenDataSoft² e a adequação dos metadados da câmara para conformidade com o DCAT-AP³. Para isso, foram realizadas diversas investigações sobre dados abertos, a sua relevância, a plataforma OpenDataSoft e o vocabulário DCAT. O método adotado incluiu a análise de documentos internos e externos pertinentes à área de atuação da câmara.

PALAVRAS-CHAVE: Gestão de dados; Dados abertos; OpenDataSoft; Câmara Municipal da Maia.

¹ <https://www.cm-maia.pt/>

² <https://www.opendatasoft.com/en/>

³ <https://op.europa.eu/en/web/eu-vocabularies/dcat-ap>

Índice de Figuras e Tabelas

Figura 1- Metodologia de Trabalho

Figura 2-Atividades de Estágio

Figura 3- Seis pilares das fases do Ciclo de Vida dos dados

Figura 4- Metadados DCAT utilizados nas plataformas

Figura 5- Catálogo de dados da Câmara Municipal da Maia

Figura 6- Metadados na Câmara da Municipal da Maia

Tabela 1-Adaptação dos Metadados conforme DCAT-AP

Tabela 2-Adaptação de mais Metadados conforme DCAT-AP

Lista de Siglas e Abreviaturas

LCI - licenciatura em Ciência da Informação

CI - Ciência da Informação

CMMaia- Câmara Municipal da Maia

UC- Unidade Curricular

UE- Regulamento de Execução

XML- Extensible Markup Language

EAD - Encoded Archival Description

CD - Catálogo de Dados

ISO- International Organization for Standardization

NUTS- Nomenclatura Unidade Territórios para Fins Estatísticos

DCAT - Data Catalog Vocabulary

GLD - Government Linked Data

DERI - Erprise Research Institute

RDF - Resource Description Framework

Agradecimentos

Gostaria de expressar o meu profundo agradecimento a todos aqueles que contribuíram para o sucesso desta experiência académica.

Primeiramente, ao Engenheiro Pedro Pimenta, desejo agradecer pelo apoio dado, orientação precisa e acompanhamento constante ao longo de todo o desenvolvimento do projeto. As suas sugestões e feedback foram cruciais para realização deste projeto.

À Câmara Municipal da Maia agradeço pela forma como fui recebida e pelos recursos fornecidos para a realização deste trabalho.

À minha tutora, Professora Mariana Malta, dedico um agradecimento especial pela sua orientação, disponibilidade e incentivo ao longo de todo o processo.

Gostaria também de estender os meus agradecimentos ao Tiago, cuja colaboração e partilha de experiências enriqueceram significativamente este trabalho.

À Susana, meu sincero agradecimento pelo apoio e encorajamento ao longo destes desafiantes anos de estudo, a sua presença e palavras motivadoras foram um verdadeiro apoio durante esta jornada académica.

Por fim, gostaria de expressar a minha profunda gratidão aos meus pais e ao meu irmão. A vossa constante presença, incentivo e paciência foram fundamentais para me manter motivada e focada ao longo deste meu percurso.

I. Introdução

Os dados representam um recurso de valor indispensável para as instituições, servindo como um pilar fundamental para a tomada de decisões, para o desenvolvimento de políticas e prestação de serviços. Neste contexto, a disponibilidade de dados abertos emerge como forma crucial para promover a transparência, permitindo que os cidadãos e outras entidades tenham acesso livre e universal a informações relevantes.

O Regulamento de Execução (UE) 2023/138 da comissão de 21 de dezembro de 2022⁴, estabelece uma lista de conjuntos de dados de elevado valor por categorias e as disposições relativas à respetiva publicação e reutilização de livre acesso. Neste sentido, o regulamento define categorias específicas e os tipos de dados que devem ser disponibilizados publicamente, tais como: geoespacial, observação da Terra e do ambiente, estatística, metrologia, empresas e propriedade de empresas, mobilidade. (Anexo I)

Além disso, no contexto europeu, a importância dos dados abertos é reconhecida através da legislação, como é exemplo nos termos do artº2 al.3, da Lei n.º26/2016 de 22 de agosto transpondo a Diretiva 2003/4/CE, do Parlamento Europeu e do Conselho, de 28 de janeiro, e a Diretiva 2003/98/CE, do Parlamento Europeu e do Conselho, de 17 de novembro⁵. Esta lei aprova o regime de acesso à informação administrativa e ambiental e de reutilização dos documentos administrativos, exigindo que as informações disponibilizadas sejam de acesso livre e universal, acessíveis, interoperáveis, íntegras, autênticas, e devidamente identificadas e localizadas.

I.1 Contexto

No âmbito da unidade curricular (UC) Projeto da Licenciatura de Ciência da Informação (LCI), foi elaborado o presente relatório que visa documentar e analisar o estágio realizado na Câmara Municipal da Maia (CMMaia) . O estágio teve início no dia 29 de abril e estendeu-se ao longo de 275 horas, com o propósito de aplicar os conhecimentos adquiridos ao longo do ciclo de estudos.

⁴ <https://eur-lex.europa.eu/legal-content/PT/TXT/PDF/?uri=CELEX:32023R0138&from=EN>

⁵ https://www.pgdlisboa.pt/leis/lei_mostra_articulado.php?nid=2591&tabela=leis&ficha=1

A Câmara Municipal da Maia localiza-se na Praça do Doutor José Vieira de Carvalho, no distrito do Porto. Esta instituição tem como fim promover o desenvolvimento do concelho e melhorar a qualidade de vida dos seus habitantes. Até ao momento, a câmara tem vindo a desenvolver o trabalho fundamental num contexto de dados abertos e, atualmente, está a desenvolver uma plataforma *OpenDataSoft* com o intuito de disponibilizar os dados à população.

I.2 Problema

O problema identificado reside na falta de alinhamento dos metadados da Câmara Municipal da Maia com os requisitos estabelecidos pelo *DACT-AP*. Esta não conformidade pode ter implicações significativas, incluindo obstáculos na interoperabilidade com outros sistemas, e dificuldades na partilha de informação.

A plataforma *OpenDataSoft* desempenha um papel crucial neste contexto, pois requer metadados consistentes e padronizados para facilitar a disponibilização e visualização acessível dos dados municipais ao público.

I.3 Objetivos do estágio

O objetivo principal deste estágio foi identificar os elementos-chave e compreender o contexto de adoção de uma plataforma de Dados Abertos em ambiente Autárquico. Em consonância com este objetivo, foram delineados os seguintes objetivos concretos: acompanhar o processo de gestão e catalogação dos conjuntos de dados, acompanhar o fluxo de dados do data lake para a plataforma *OpenDataSoft*, adequação dos metadados com o padrão DCAT-AP e a plataforma *OpenDataSoft* e colaborar na monitorização da utilização da plataforma.

I.4 Organização do relatório

Este relatório está dividido em seis seções e os anexos. Após a introdução, faz-se uma breve contextualização sobre ciência da informação, seguida pelo estado da arte de dados abertos, metadados, catálogo de dados, *data lake*, ISO 37120, *OpenDataSoft* e *DCAT*. Posteriormente, apresenta-se a concretização do projeto e a conceção da solução. O relatório termina com uma análise dos resultados e uma reflexão crítica do projeto.

2. Metodologia e plano de trabalho

O estágio teve início no dia 29 de abril até ao dia 28 de junho. Foi adotado um idem de trabalho híbrido, o que possibilitou a realização de tarefas tanto na Câmara como remotamente, a partir de casa. Esta abordagem flexível permitiu adaptar o trabalho às necessidades individuais e às exigências específicas de cada tarefa. Durante o estágio, foram identificadas quatro fases essenciais de desenvolvimento.



Figura 1- Metodologia de Trabalho

Para garantir uma organização eficiente e fácil das tarefas, foi desenvolvido um cronograma, no qual facilitou a monitorização do progresso e a gestão do tempo, assegurando que todas as tarefas fossem realizadas dentro dos prazos estipulados.



Figura 2-Atividades de Estágio

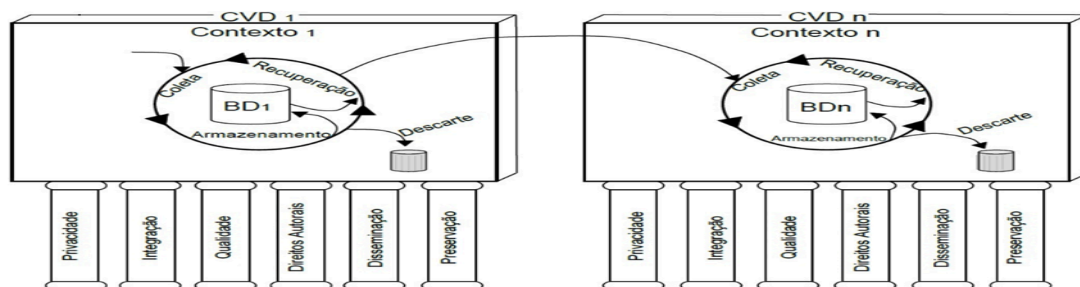
3. Contextualização em Ciência da Informação

Ciência da informação (CI) pode ser compreendida como um campo de estudo, tendo como foco principal a informação estudada pelo ser humano. Abrange a interpretação, organização, transformação, armazenamento e a recuperação, tanto no uso quanto na transmissão da informação. Considera-se a CI como transdisciplinar e, logicamente, possui um perfil interdisciplinar, o que a coloca efetivamente no campo das Ciências Sociais e em relações disciplinares intensas com as Ciências Exatas e Naturais. Nesse sentido, o estudo do fluxo infocomunicacional abrange a origem, a organização e o comportamento informacional.

Na literatura da CI, encontra-se uma diversidade de vocabulários, sendo o dado um termo central de investigação. A CI desempenha um papel crucial no desenvolvimento, acesso e utilização dos dados, reconhecendo a importância da sua origem. Nesse sentido, a aplicação do Ciclo de Vida dos Dados (CVD) é essencial no contexto da ciência da informação. O CVD pode ser considerado uma curadoria em ação, estão envolvidas quatro fases. Numa primeira fase, a coleta, o objetivo é compreender a necessidade informacional e o seu contexto de uso dos dados, realizando uma análise dos dados. Na segunda fase, o armazenamento, o foco é identificar que dados são armazenados, para que fins, em que suportes, e garantir sua acessibilidade.

Numa Terceira fase, recuperação, concentra-se no acesso e uso dos dados, tomada de decisão se os dados são mantidos e a viabilização da recuperação dos dados. Numa última fase, o descarte, decide-se se os dados serão eliminados ou preservados, considerando a sua utilidade para usos futuros. Em todas estas fases estão presentes seis pilares: privacidade, integração, qualidade, direitos autorais, disseminação e preservação.

Figura 3- Seis pilares das fases do Ciclo de Vida dos dados



Fonte-Adaptado de Sant'Ana (2016)

No universo dos dados, nem tudo pode ser tratado sem autorização, uma vez que existem direitos sobre estes. Quando os dados contidos em dados públicos incluem informações pessoais ou confidenciais, é fulcral saber como gerenciá-los de forma correta, sem comprometer a segurança e privacidade das pessoas. Assim, o profissional da informação desempenha um papel fulcral nesta área, pois é responsável por identificar que tipo de dados são e avaliar sua confiabilidade, integridade e conformidade considerando os mecanismos de controle e adotar práticas para salvaguardar os dados sensíveis.

4. Estado da Arte

Este ponto faz referência a toda informação e conceitos que irá suportar o trabalho apresentado.

4.1 Dados/ Dados Abertos

No contexto atual, a valorização do dado transcende a mera acumulação de informações, assumindo uma relevância vital para as organizações em diversos setores. O dado pode ser considerado a menor unidade de informação, que isoladamente não tem valor significativo até ser contextualizado e interpretado. Quando organizados e tratados, os dados transmitem informação. Dados combinados geram informação sobre algo e são considerados informações quando quantificados ou quantificáveis.

O dado é um ativo estratégico, não apenas ressalta sua importância intrínseca, mas também destaca a necessidade premente de sua adequada manipulação e disseminação para que alcance seu potencial máximo como fonte de informação.

Assim, é imprescindível compreender que os dados, por si só, não se transformam em informação útil sem o devido tratamento, categorização e contextualização. Os dados Abertos são informações que podem ser livremente usadas, reutilizadas e redistribuídas por qualquer pessoa. Para que os dados sejam considerados abertos, devem atender a certos critérios, incluindo a disponibilização sob uma licença livre, que permite o uso sem limitações de direitos autorais e de controle.

Além disso, sob uma ótica ética e legal, os dados devem ser disponibilizados exclusivamente sob licenças abertas, assegurando sua integridade e preservando sua fidelidade em relação à fonte original. A formatação dos dados também desempenha um papel crucial na sua acessibilidade e usabilidade. Optar por formatos abertos, que não imponham barreiras à sua utilização, é essencial para garantir que a informação seja facilmente interpretada.

Nos termos de dados abertos, uma *Application Programming Interface* (API) é um software que tem como objetivo proporcionar o acesso direto de um sistema aos dados de outro sistema. Uma *API* elimina a necessidade de baixar conjuntos de dados internos, permite que o conjunto de dados seja automatizado e permite que os dados estejam em variadíssimas

ferramentas. Uma *API* é fundamental quando os dados são atualizados, acessíveis diretamente e utilizados/reutilizados por terceiro.

A importância dos dados para câmara Municipal é inegável, pois este desempenha um papel crucial na tomada de decisões com critério, ajuda na gestão eficiente do território, ajuda na identificação de padrões e na otimização de recursos.

Os dados sobre criminalidade, por exemplo, permitem à câmara analisar padrões e identificar áreas com maior risco de crimes. Esta análise facilita a implementação de medidas preventivas e de segurança específicas. Já, na área da educação, a câmara pode utilizar dados sobre o número de crianças matriculadas e o desempenho académico dos jovens para planear e melhorar a oferta educativa. Assim, ajuda prevenir as necessidades da população e promove uma educação de qualidade. Além disso, a análise dos dados de consumo, como o consumo de água e energia, pode ajudar a câmara a gerir melhor os recursos. Compreender os padrões de consumo permite implementar políticas de eficiência e sustentabilidade, garantindo que os recursos são utilizados de forma consciente e responsável, o que é fundamental para a preservação do ambiente e a redução de custos.

4.2 Metadados

Os metadados desempenham um papel crucial na gestão e análise de dados. Apresentam informações sobre os dados, facilitando a organização, localização, identificação, categorização e recuperação eficiente das informações. Em essência, os metadados são "dados sobre dados" que descrevem outros dados. Segundo Ortiz-Repiso Jiménez (1999, p. 218), os metadados são definidos como: "Um conjunto de dados que pode ser usado para descrever e representar recursos informacionais."

A normalização dos sistemas descritivos de informação é fundamental, pois implica a adoção de padrões que asseguram a consistência e a interoperabilidade e de diferentes plataformas e sistemas. Para isso, existe uma variedade de padrões de metadados que podem ser aplicados em estruturas informacionais, desde os mais simples até os mais complexos. Os metadados simples, por exemplo, são aqueles não estruturados e são extraídos automaticamente, sem a necessidade de intervenção humana. Por outro lado, os formatos estruturados existem a participação ativa de profissionais da informação para sua criação e manutenção.

Os Metadados descritivos, fornecem informações detalhadas e identificam os recursos informacionais, podendo incluir elementos como título, autor, assunto, resumo e identificador, muitas vezes utilizando linguagens como *HTML* e vocabulários controlados para sua representação.

Já, os metadados estruturais, são compostos por informações que o documento possui, nos quais diversos elementos devem ser recompostos e organizados, visando facilitar a navegação e a apresentação dos dados, empregando linguagens específicas como *Extensible Markup Language* (XML) e *Encoded Archival Description* (EAD).

Por outro lado, os metadados Administrativos, fornecem informações sobre quando, como e porquê de o recurso foi criado, também facilita a gestão e o processamento dos repositórios digitais (MOA2 e OAIS).

Deste modo, a utilização adequada e a padronização dos metadados revelam-se cruciais para a eficácia dos Sistemas de Informação, contribuindo para uma melhor gestão e aproveitamento dos recursos informacionais disponíveis.

4.3 Data Lake

O conceito *data Lake* foi introduzido por James Dixon no ano de 2010. Consiste num repositório de armazenamento que contém grandes volumes de dados, independentemente de estarem estruturados ou não, retirados de seu formato original. Manter a ordem de chegada dos dados é fundamental para preservar sua integridade. Além disso, a acessibilidade e a agilidade são aspetos essenciais deste conceito, e uma única ingestão de dados é crucial para garantir a para preservar a consistência e confiabilidade dos dados, evitando duplicatas.

A implementação de um data lake pode variar de acordo com a área de negócio, os dados e os objetivos da instituição e prometer reduzir custos de processamento e armazenamento, no entanto, há etapas gerais que devem ser seguidas, como reunir dados originais das fontes e a identificação da arquitetura a ser usada. As abordagens para um data Lake podem envolver a combinação de diferentes tipos de bancos de dados, como *SQL* e *NoSQL*, além de técnicas de processamento analítico online (OLAP). Para concluir, *Sawadogo e Darmont (2021)* chegaram à conclusão de que um data Lake é simplesmente um sistema escalável de armazenamento e análise de dados de qualquer tipo, mantidos no seu formato original e utilizado principalmente por especialistas em dados para extrair conhecimento.

A Câmara Municipal da Maia criou um “*data lake*”, uma base de dados que permite armazenar um conjunto de dados como dados geográficos. Existe também uma Base de dados de séries temporais, que deriva de várias fontes, como instituto nacional de estatística, direção geral da Energia e Geologia, entre outros. Todos os dados estão em formato *JavaScript Object Notation (JSON)*. É o formato de dados mais utilizado para comunicação assíncrona entre o navegador e o servidor. A utilização deste formato apresenta, semanticamente, o modelo e estrutura de dados leves e fáceis de ler. A simplicidade, legibilidade, interpretação por uma variedade de linguagens, programação e a fácil integração com *APIS* são características deste formato.

4.4 Catálogo de Dados

Um Catálogo de Dados (CD) é uma ferramenta útil no contexto de sistemas analíticos, permitindo uma gestão eficiente dos metadados associados aos conjuntos de dados. Com o CD, é possível reunir, descrever e organizar de maneira lógica as informações mais importantes sobre os dados, incluindo seu conteúdo, estrutura, origem, formato entre outros. A ideia principal dos CD é que os utilizadores acedam aos dados de uma forma eficaz. Utilizando metadados, o CD serve como um inventário abrangente, oferecendo uma visão única e completa dos dados disponíveis.

4.5 ISO 37120

A norma da *International Organization for Standardization (ISO)*, criada em 1946 com o propósito de promover a normalização de produtos e serviços através da utilização de normas internacionais, visa a melhoria contínua em diversas áreas. A ISO 37120 foi a primeira norma internacional dedicada à sustentabilidade urbana, focada na formação de cidades inteligentes. Essencial para Câmara Municipal, esta norma é composta por indicadores de serviços e qualidade de vida, e permite uma gestão e tomada de decisões fundamentadas na análise de dados. O planeamento e desenvolvimento sustentável das cidades, bem como a avaliação de impactos concretos dos projetos realizados, são apresentados por esta norma.

Um indicador é uma ferramenta que ajuda obter informações, medir o desempenho de gestão de serviços, a qualidade de vida e o desenvolvimento de políticas da cidade e permite a comparação. Os indicadores da ISO 37120:2018 são indicadores de sustentabilidade, essenciais

para monitorizar o desenvolvimento sustentável da cidade, trazem métricas percentuais e são definidos por temas: economia, educação, energia, meio ambiente e clima, finanças, governo, saúde, habitação, população e condições sociais, lazer, segurança, resíduos sólidos, desporto e cultura, telecomunicações, transporte, agricultura urbana/local e segurança alimentar, planeamento urbano, águas residuais e água. Estes indicadores são cruciais para auxiliar as cidades, medir o progresso, medir a qualidade de vida dos seus habitantes, permitindo comparações de melhoria e apoiando no desenvolvimento de políticas públicas e estabelecimento de prioridades.

4.6 OpenDataSoft

A plataforma que a Câmara Municipal Maia utiliza designa-se *OpenDataSoft*. Esta plataforma pertence a uma empresa francesa fundada em 2011, que se destaca no campo do compartilhamento de dados. Através do seu serviço especializado, oferece uma solução completa para organizar, partilhar, visualizar e reutilizar informações de forma eficiente e acessível.

OpenDataSoft é uma plataforma de dados em nuvem que permite publicar e gerenciar os dados de forma fácil, segura e através de uma interface, não se limita apenas a instituições públicas, mas também atende às necessidades de empresas privadas. Os utilizadores podem reunir os dados manualmente ou de forma automática, fornecendo um *URL* que permite o acesso direto às informações desejadas. Esta abordagem simplificada torna a manipulação e a partilha de dados uma tarefa acessível a uma variedade de utilizadores e permite visualização interativa de dados, gráficos e mapas, desde administradores públicos até gestores de empresas privadas. Além disso, a *OpenDataSoft* disponibiliza cinco *APIs* diferentes, cada uma com funcionalidades específicas:

- *ODS Explore API V2*: A *API* principal da *OpenDataSoft* é a *Explore*, proporciona acesso a dados públicos na plataforma. Esta *API* permite a solicitação de registos, disponibiliza os dados em formato *JSON*. É possível exportar o conjunto de dados e analisá-los.
- *Automation API*: Permite automatizar os processos de governança de dados e garante que os dados estejam atualizados.
- *WFS e CSW*: São padrões que se centram em dados geográficos.

Adopção da plataforma OpenDataSoft para publicação de dados abertos

- *aOData*: Um padrão para APIs Rest que disponibiliza uma linguagem para executar solicitações em APIs. No entanto, possui a desvantagem de ser bastante complexo.
- *riple Pattern Fragments API*: Permite consultas de padrão triplo em conjuntos de dados de portais *OpenDataSoft* e facilita consulta de dados no formato RDF.

Estas APIs podem ser usadas para diversas finalidades, incluindo pesquisas de conjunto de dados, análises detalhadas e execuções de agregações geográficas. A escolha da *API* adequada depende do contexto específico e dos objetivos do projeto.

A *OpenDataSoft* melhora a tomada de decisões permitindo ter acesso aos dados reais, aumenta a inovação, melhora na colaboração permitindo que os funcionários tenham acesso à plataforma e aumenta a transparência.

Em 2014, a *OpenDataSoft* foi reconhecida pelo *Electronic Business Group* como uma das dez start-ups mais inovadoras da França, demonstrando seu compromisso com a inovação e excelência no campo da tecnologia da informação.

4.7 DCAT

O *DCAT* é uma descrição baseada no Vocabulário de Catálogo de Dados (*DCAT*). Inicialmente, o *DCAT* foi criado pelo *Enterprise Research Institute* (*DERI*), tendo sido redefinido pelo *Gov Interest Group* e padronizado em 2014 pelo grupo de trabalho *Government Linked Data* (*GLD*). A versão mais recente do *DCAT*, a versão 3, foi desenvolvida pelo *Dataset Exchange Working Group*.

O *DCAP* é um perfil de aplicação *RDF* (*Resource Description Framework*), permite definir formalmente a semântica dos metadados e saber o significado dos elementos de metadados, este modelo é processável por máquinas computacionais. No entanto, também utiliza de termos de outros vocabulários na normalização dos dados:

- *Dublin Core Metadada Initiative Metadata Terms* (*DC terms*)
- *Friend of a Friend Vocabulary* (*FOAF Vocabulary*)
- *vCard Ontology*
- *PROV Ontology* (*PROV-O*)
- *Vocabulary of Interlinked Datasets* (*VOID*)

A utilização de um padrão de vocabulário facilita a descoberta de conjuntos de dados. Os dados apresentados num catálogo podem estar disponíveis em diversos formatos, tais

como *XML* e *RDF*. A interoperabilidade entre os formatos é crucial e conduz à existência de seis classes principais nas quais o *DCAT* se baseia: (Anexo 2- constituição de metadados obrigatórios por classe)

- *dcat: Catalog*- ilustra um catálogo, que engloba um conjunto de dados, onde cada item é um registo de metadados que descreve um recurso específico.
- *dcat: Resource*- ilustra uma coleção de dados, um serviço de dados ou outro recurso com objetivo de ser descrito por um registo de metadados em um catálogo. Esta classe é considerada a classe pai dcat: Dataset de dcat: DataService e dcat:Catalog .
- *dcat: Dataset*- ilustra um conjunto de dados. Uma coleção de dados publicidades é um conjunto de dados, embora, muitas das vezes, os dados não estejam em um único formato.
- *dcat: Distribution*- ilustra uma forma simples de um arquivo para download onde os dados são fáceis de aceder.
- *dcat: DataService*- representa um serviço de dados, incluindo uma coleção de operações utilizando uma interface (API).
- *dcat: CatalogRecord*- ilustra um elemento de metadados no catálogo, especialmente relacionado com os dados de registo, como o responsável pela inclusão do elemento e a data de inserção.

No que diz respeito à segurança e privacidade, o *DCAT* pode conter dados com informação pessoal ou privada, para além de os metadados podem conter informações sobre criadores e editores. Por isso, devem ser implementadas medidas para garantir a segurança e privacidade. De igual modo, os dados sensíveis devem ser armazenados de forma segura, de acordo com os requisitos legais e a integridade e autenticidade dos dados devem ser preservados. Além disso, é necessário assegurar a integridade da *API*.

O perfil de aplicação *DCAT* para portais de dados na Europa - *DCAT-AP* - possibilita uma especificação comum para ilustrar um conjunto de dados no setor público na Europa. Esta padronização ajuda na partilha de descrições de conjuntos de dados entre diferentes tipos de portais de dados. Com *DCAT-AP* permite ter um catálogo de dados que proporciona uma descrição uniforme de um conjunto de dados, mantendo o seu próprio sistema para documentar, armazenar e agregar conteúdos, como é o caso do Portal Europeu de Dados, com o objetivo de ter um único ponto de acesso. Assim, os objetivos do *DCAT-AP* são facilitar

a descoberta e o acesso a conjuntos de dados do setor público na Europa, promover a interoperabilidade entre portais de dados, melhorar a qualidade dos metadados dos conjuntos de dados e aumentar a transparência e a responsabilidade do governo.

Uma das dificuldades na implementação do *DCAT-AP* reside em saber gerir a qualidade dos metadados. Por exemplo, na conformidade dos metadados, é primordial garantir que os elementos tenham o mesmo tipo de dados prescritos e que os valores reais nos metadados sejam precisos, representando adequadamente o conjunto de dados e sua natureza.

Para garantir o sucesso da implantação do *DCAT-AP* é recomendado analisar a conformidade dos metadados com o *DCAT-AP*, considerando elementos obrigatórios, tipos de dados e vocabulário controlados. Assim, para verificar se os metadados estão em conformidade, recomenda-se utilizar o validador *DCAT-AP* desenvolvido pela Comissão Europeia. Outra alternativa pode ser a utilização do shapes constraint *language- SHACL do W3C*, uma linguagem para validar grafos *RDF*.

Para concluir, o *DCAT* desempenha um papel crucial como um vocabulário para a publicação de dados em catálogos. Ao proporcionar uma estrutura padronizada para descrever conjuntos de dados, oferece uma abordagem consistente e organizada para a divulgação de informações. A sua utilização não apenas facilita a disponibilização de dados ao público, mas também promove a interoperabilidade e a compreensão entre diferentes plataformas e sistemas. Portanto, considerar o *DCAT* como uma escolha para padronizar a descrição de dados é essencial para promover a transparência, acessibilidade e eficácia na disseminação de informações.

GeoDCAT-AP

O *GeoDCAT-AP* representa uma ampliação do *DCAT-AP*, destinada a descrever conjuntos de dados geoespaciais e séries de conjuntos, com uma ligação de sintaxe *RDF* para união de elementos de metadados estabelecidos na ISO 19115:2003 e os determinados na diretiva INSPIRE. Importa salientar que as especificações *GeoDCAT-AP* não alteram o Regulamento de Metadados INSPIRE, nem as suas diretrizes técnicas.

A ISO 19115:2003⁶, substituída pela ISO 19115:2014⁷, disponibiliza um modelo para ilustrar as informações geográficas, define os elementos de metadados obrigatórios, opcionais

⁶ <https://www.iso.org/standard/26020.html>

⁷ <https://www.iso.org/standard/53798.html>

e as suas propriedades. Esta norma pode ser aplicada na catalogação de vários recursos, atividades de câmara, serviços geográficos e conjuntos de dados geográficos e séries de conjuntos de dados.

A diretiva INSPIRE entrou em vigor em 15 de maio de 2007⁸ e tem como fim originar uma Infraestrutura de Dados Espaciais da UE, aborda 34 temas relacionados com os dados espaciais. Os Princípios INSPIRE⁹ baseiam-se que os dados só devem ser recolhidos uma única vez, a disponibilidade da informação deve ser transparente fácil e será possível combinar informações de diferentes fontes na Europa. O principal objetivo *GeoDCAT-AP* é a partilha de informações entre instituições do setor público de modo fácil e reutilizável.

O *GeoDCAT-AP* tem que incluir os componentes obrigatórios do *DCAT-AP*, ou seja, toda a descrição de metadados válida em *DCAT-AP* deve ser válida em *GeoDCAT-AP*. (Anexo 3 apresenta elementos do *GeoDCAT-AP*)

O *GeoDCAT-AP* apresenta todos os elementos de metadados do perfil da ISO 19115 e os definidos no INSPIRE, porém, os componentes opcionais podem ser excluídos. O *GeoDCAT-AP*, permite uma maior visibilidade e usabilidade e ajuda na interoperabilidade. (Anexo 3 ilustra os metadados INSPIRE e os metadados Geoespaciais da ISO 19115:)

StatDcat-AP

O *StatDCAT-AP* é uma extensão do *DCAT-AP* para ilustrar conjuntos de dados referentes ao domínio estatístico. O perfil de aplicação *StatDCAT* foi desenvolvido em 2015 por um grupo de trabalho *StatDCAT-AP*, que inclui representantes do *Eurostat* e do Serviço de Publicações da UE, composto por membros das instituições da UE.

O *StatDCAT-AP* define um conjunto de metadados específicos para descrever dados derivados de estatísticas. O principal objetivo é facilitar a descoberta de novos conjuntos de dados, fornecer uma especificação compatível com *DCAT-AP* e melhorar a interoperabilidade entre as descrições de dados estatísticos.

O modelo *StatDCAT-AP* abrange as quatro entidades principais do *DCAT-AP*: catalogue, catalogue record, dataset, distribution. Foram criados novos requisitos para

⁸ <https://eur-lex.europa.eu/legal-content/PT/TXT/?uri=celex%3A32007L0002>

⁹ https://knowledge-base.inspire.ec.europa.eu/index_en

complementar este novo perfil. (Anexo 4 ilustra alguns exemplos dos metadados *StatDACT-AP*)

4.8 OpenDataSoft e DCAT

A plataforma *OpenDataSoft* disponibiliza um modelo pré-definido em *OpenDataSoft* valida automaticamente os metadados do conjunto de dados para garantir que estejam em conformidade com o *DCAT-AP*. Adicionalmente, é possível exportar os dados no formato DCAT-AP para publicação em portais que seguem este padrão.

As *APIs* na *OpenDataSoft* possibilitam a obtenção de dados remotamente através do protocolo *HTTP*. Uma nova *API* é criada sempre que publica um novo conjunto de dados. Estas *APIs* são fundamentais para a integração de sistemas e o consumo de dados por diversas aplicações, oferecendo um meio eficiente e automatizado de acesso aos dados. Esta abordagem facilita a harmonização e interconexão de sistemas, otimizando processos de trabalho e colaboração.

Figura 4- Metadados DCAT utilizados nas plataformas

DCAT	CKAN	Socrata	OpenDataSoft
dcat:Dataset → dct:title → dct:description → dct:issued → dct:modified → dct:identifier → dcat:keyword → dct:language → dct:publisher → dct:contactPoint → dct:accrualPeriodicity → dct:landingPage → dct:theme	title notes metadata_created metadata_modified id tags language organization maintainer, author (-email) frequency url -	name description createdAt viewLastModified id tags owner tableAuthor - - - category	title description - modified datasetid keyword language publisher - - - theme
dcat:Distribution → dct:title → dct:issued → dct:modified → dct:license → dcat:accessURL → dcat:downloadURL → dct:format → dct:mediaType → dct:byteSize	resources.name resources.created resources.last_modified license_{id, title, url} resources.url resources.download_url resources.format resources.mimetype resources.size	- - - licenseId export URL ^a - export format ^a export mime-type ^a -	- - - license export URL ^a - export format ^a export mime-type ^a -

Fonte: *Automated Quality Assessment of Metadata across Open Data Portals*

Conforme mostrado na figura acima, apresenta chaves de metadados de diferentes plataformas. Os campos vazios na tabela indicam que não há correspondência com uma chave de mesmo significado semântico. No entanto, é evidente que o mapeamento de chaves dos metadados *OpenDataSoft* utiliza os mesmos nomes de chaves do vocabulário *DCAT*. Este mapeamento é essencial para garantir que todos os elementos de dados relevantes sejam traduzidos corretamente entre os dois sistemas, preservando a integridade e a utilidade das informações.

5. Concretização do projeto

Nesta seção, são descritas as opções de implementação do projeto e as condições de aplicação dos instrumentos desenvolvidos. O principal objetivo desta fase é detalhar o processo de adequação dos metadados da Câmara Municipal da Maia ao padrão *DCAT-AP*.

O projeto principal concentrou-se na utilização dos dados abertos disponíveis no catálogo da Câmara Municipal da Maia, com foco na padronização dos metadados para conformidade *DACT-AP*. O problema inicial identificado foi a necessidade de estruturar os dados de forma acessível e útil, permitindo análises eficientes e aplicação em plataformas como o *OpenDataSoft*.

Além disso, numa fase inicial foram realizadas várias tarefas na plataforma *OpenDataSoft*, incluindo a inserção dos *links* do *Power BI* nos *datasets*, a verificação ou correção dos títulos, a verificação dos temas atribuídos, a ordenação das tabelas por ordem crescente, desativar visualização *standard* e a verificação dos temas *ODS ICONS*. (Anexo 5)

Para concretização deste projeto foram seguidos os seguintes passos: levantamento de requisitos, ou seja, identificação dos requisitos específicos do *DCAT-AP*, avaliação dos metadados atuais para identificar discrepâncias, conversão e alinhamento dos metadados existentes com o padrão *DCAT-AP*.

4.1 Desenho da solução, apresentação, avaliação e discussão dos resultados

A solução proposta para resolver o problema identificado, garante a conformidade dos metadados da Câmara Municipal da Maia com o *DACT-AP* e assegura a qualidade e a integridade dos dados públicos disponibilizados. Este processo será guiado por diretrizes claras baseadas nos requisitos do *DACT-AP* e nas melhores práticas de gestão de metadados.

Análise ao metadados

A Câmara Municipal da Maia elaborou um catálogo de dados com um total de 16 indicadores. (Anexo 6- realiza-se uma análise dos metadados atuais da câmara)

A ficha técnica de cada indicador apresenta a seguinte estrutura:

Lista de indicadores

1. Ação Social
2. Administração Pública e Serviços
3. Ambiente
4. Cultura e Património
5. Economia
6. Educação
7. Energia
8. Habitação
9. Lazer e Desporto
10. Mobilidade
11. Planeamento Urbano
12. População
13. Saúde
14. Segurança
15. Outros
- 16.

Ficha técnica

Sigla/home curto	CoeGm
Indicador ^a	Coefficiente de Gini do rendimento bruto declarado deduzido do IRS liquidado por sujeito passivo
Descrição Detalhada	Coefficiente de Gini do rendimento bruto declarado deduzido do IRS liquidado por sujeito passivo (em %) - O coeficiente de Gini é um indicad
Tema	Ação Social ▼
Forma de Cálculo	Não tem fórmula disponível no INE
P	tiny ↗
Fonte [Ⓢ]	INE
Origem [Ⓢ]	Estatísticas do rendimento obase na info. prod. pelo Min. das Finanças-Aut.Tributária e Aduaneira.
Unidades territoriais [Concepto] ▼	Unidade(s) [t]
Forma de actualização [BEST API] ▼	Periodicidade de Actualização [anual] ▼
Observações	
P	
Editor	soufia.maia@cm-maior.pt ▼

Coeficiente de Gini do rendimento bruto declarado deduzido do IRS liquidado por sujeito passivo

Fonte: INE Coeficiente de Gini do rendimento bruto declarado deduzido do IRS liquidado por sujeito passivo (em %) - O coeficiente de Gini é um indicador representativo da desigualdade na distribuição dos rendimentos que varia entre zero (valor a ausência dessa distribuição. Assume valores entre 0 (quando todos os indivíduos têm rendimento igual) e 100 (quando todo o rendimento se concentra num único indivíduo)

Ano	Coeficiente de Gini (%)
2015	~82
2016	~78
2017	~72
2018	~65
2019	~35
2020	~35
2021	~35
2022	~35
2023	~35
2024	~35

Origem: Estatísticas do rendimento obase na Info. prod. pelo Min. das Finanças- Aut.Tributária e Aduaneira. | Fonte: INE (Última actualização: 2023-12-18)
Mais info: <https://www.inec.pt/pt/indicadores/estatisticas-de-rendimento-obase-na-info-prod-pelo-min-das-financas-aut-tributar-e-aduaneira/>

Incluído no dataset: 2023-05 (Última actualização: 2023-12-18) (Última tentativa de actualização: 2023-12-18 05:00:00)

Notas: Dados referidos em "SGO720 na Câmara Municipal de Maior" (Winning, 2021, pág. 20). Os dados do INE referem-se ao período 2020.

Editor/Curador: BIF (Bureau figureado/cm-maior.pt)

Contact point: Maia - Portal de dados | Contact email: sugestao@bif.cm-maior.pt

Licença: Creative Commons CCZero ©Creative Commons, CC-BY

Figura 5- Catálogo de dados da Câmara Municipal da Maia

No entanto, alguns metadados não estão de acordo com o DCAT-AP. Para assegurar a conformidade com este padrão, são necessárias as seguintes alterações:

O termo "sigla/nome curto" será substituído por "título"(*title*), que corresponde essencialmente ao nome do recurso.

○ termo “indicador” passa a ser designado “identificador”(“*identifier*”), permitindo atribuir um código único ao recurso que está a ser descrito.

No DCAT-AP, não existe o termo "descrição detalhada". Apenas o termo "descrição" (*"description"*) está disponível, que consiste num texto livre para descrever o recurso.

Por outro lado, no contexto do DCAT-AP, não foi identificado nenhum termo adequado que corresponda à descrição de "Forma de cálculo". Portanto, não há um termo específico disponível para descrever essa funcionalidade dentro das diretrizes do DCAT-AP. Deste modo, a fórmula de cálculo poderia ser colocada num comentário.

A palavra "fonte" não está presente no vocabulário do *DCAT-AP*. Existe a possibilidade que pode ser adotado o termo “referenciado por” (“referencedy”) no qual permite que um recurso se relacione com uma publicação que faz referência ou cita um conjunto de dados. Este termo é adequado para situações em que se deseja indicar que os dados foram utilizados

em artigos, relatórios ou outros documentos. Embora, embora seja relevante para indicar que há outros recursos (como publicações ou citações) que fazem referência ao recurso catalogado, não descreve diretamente a fonte de onde o recurso pode ser.

No entanto, recentemente em junho de 2024 foi publicado o DCAT-AP 3, no qual surgiu o termo “*source*” (“fonte”), que relaciona o conjunto de dados descritos com a sua derivação.¹⁰

Já o termo “origem” passa a ser “criador” (“*creator*”). Esta categoria designa a entidade responsável pela criação original dos dados, identificando a fonte primária ou instituição que gerou as informações. Isto proporciona clareza sobre a autoria e a responsabilidade pela produção dos dados disponibilizados.

Embora o termo “Observações” não esteja presente diretamente no *DCAT-AP*, o termo “comentário” (“*comment*”) pode ser uma boa substituição, uma vez que cumpre a função de adicionar comentários adicionais ou fornecer informações contextuais importantes sobre os recursos. No *DCAT-AP*, há um campo chamado `dct:description` que pode ser utilizado para adicionar descrições e comentários adicionais sobre os recursos. Outra possibilidade é usar `dct:comments` ou `rdfs:comment`, dependendo do contexto e da estrutura de dados que estão sendo mapeados.

Português	Inglês	Explicação
Título	title	Nome dado ao recurso
Identificador	identifier	Identificador exclusivo do recurso descrito
Descrição	description	Texto sobre o recurso
Tema	theme	Tema que se refere
Referenciado por	referencedby	Recurso relacionado/ faz referência ou cita
Criador	creator	Entidade responsável pela produção do recurso
Comentário	Comment	Comentário adicional
Editor	Publisher	Entidade responsável por

¹⁰ <https://semiceu.github.io/DCAT-AP/releases/3.0.0/#CatalogueRecord>

		disponibilizar o recurso
--	--	--------------------------

Tabela I-Adaptação dos Metadados conforme DCAT-AP

Existem ainda mais metadados, que são os seguintes:

```
"sql": "selecione ano, 'CoeGini' de base2IRA ordenar por ano",
"q": "[2015, 2016, 2017, 2018, 2019, 2020, 2021, 2022, 2023, 2024],
"q": [30,2, 37,4, 36,7, 35,9, 35,2, 35,1, mulo, mulo, mulo, mulo],
"metadados": {
  "nome": "CoeGini",
  "descricao": "Coeficiente de Gini do rendimento bruto declarado deduzido do IRS liquidado por sujeito passivo",
  "descrit": "Coeficiente de Gini do rendimento bruto declarado deduzido do IRS liquidado por sujeito passivo (em %). O coeficiente de Gini é um indicador representativo da desigualdade na distribuição dos rendimentos que visa sintetizar num único valor a assimetria dessa distribuição. Suponha entre valores entre 0 (quando todos os indivíduos têm rendimento igual) e 100 (quando todo o rendimento se concentra num nico indivíduo)",
  "origem": "Estatísticas de rendimento c/base na info. prod. pelo Min. das Finanças- Aut.Tributária e Aduaneira.",
  "fonte": "INE",
  "MetaInfUrl": "https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_indicadores&indOcorrCod=0009943&selTab=tab0",
  "PrimoPref": "",
  "UltimoPref": "2021",
  "DataUltimaActual": "2023-12-18",
  "DataReg": "06/03/2023 14:16:14",
  "DataUltimaActualLocal": "2023-12-18 11:00:00",
  "DataUltimaVerifica": "2023-12-18 11:00:00",
  "Tema": "Acção Social",
  "FormaCalculo": "Não tem fórmula disponível no INE",
  "DisTribucional": "Concelho",
  "Licenca": "Licença: Creative Commons CCZero",
  "LicenseURL": "http://opendefinition.org/licenses/cc-zero/",
  "Editor": "BF (bruno.figueiredo@cm-maia.pt)",
  "ContactPerson": "Maia - Portal de dados",
  "ContactEmail": "suporte.dados@cm-maia.pt",
  "com": "Dados referidos em 'a href='http://bit.ly/38xy09' targets='\">ISO37120 na Câmara Municipal da Maia/as' (Winning, 2021, pág. 205) . Os dados do INE referem-se ao período de 2021."
},
"fonte": {
  "apiurl": "https://base.cm-maia.pt/Ba2e/api/ap14s.php",
  "exemplo": "https://base.cm-maia.pt/Ba2e/api/ap14s.php?nome=habit",
  "JSON validado em": ["https://www.itb.ec.europa.eu/json/any/upload", "https://jsonlint.com/"]
}
```

Figura 6- Metadados na Câmara da Municipal da Maia

O termo "*nome*" é substituído por "*título*" ("*title*"), que representa o nome dado ao recurso.

"PrimoPref" passa a ser denominado "startDate", indicando o início do período (com o domínio de "*periodOfTime*" que ilustra um intervalo de tempo definido) .

"UltimoPref" passa a ser "*endDate*", referindo-se ao fim do período de dados (com o domínio de "*PeriodOfTime*"). Adicionalmente, é ainda possível considerar a inclusão do mês e data completa em etapas subsequentes.

“*Year*” para representar apenas o ano ("2024"),

“*YearMonth*” para representar ano e mês ("2024-08"),

“*date*” para representar uma data completa (“2024-08-30”),

“*dateTime*” para representar uma data e hora completa ("2024-08-30 14:30:00").

O termo "*DataUltimaActual*" será renomeado para "*modified*", indicando a data mais recente em que o recurso foi alterado, atualizado ou modificado. Esta propriedade refere-se a mudanças no próprio recurso, não no registro do catálogo.

Já o termo "*RegDate*" (data de registo) será substituído por "*issued*", representando a data formal de emissão da publicação do recurso. Esta propriedade deve ser usada para indicar a primeira data de emissão conhecida.

Por outro lado, a "*DataUltimaActuaLocal*" também será alterada para "*modified*", indicando novamente a data mais recente em que o recurso foi alterado, atualizado ou modificado. E o termo "*DataUltimaVerfica*" será renomeada para "*modified*", representando a data da última verificação dos dados.

O termo "*DivTerritorial*" será substituído por "*spatial*", que se refere à área geográfica coberta pelos conjuntos de dados descritos. Para especificar uma região espacial, pode-se utilizar o termo "*Location*".

Observa-se que o termo licença ("*license*") está conforme o *DCAT-AP*. No entanto, caso se deseje especificar o padrão ao qual o recurso descrito está em conformidade, pode-se utilizar o termo "*conformsTo*".

"Editor" está correto, mas pode gerar ambiguidade, uma vez que em inglês a palavra correta é "publisher", referindo-se à entidade responsável por disponibilizar o recurso. Além disso, para complementar este recurso, recomenda-se os termos "*Agent*" para quando para quando é utilizado um pessoa "*foaf:Person*" é uma organização"- "*foaf: Organization*".

Por outro lado, o "*ContactoPerson*" seria alteraria para a "*contact Point*", que contém informações de contacto sobre a quem consultar em caso de dúvidas sobre os dados

Neste ponto, recomenda-se o uso do vard:

"@type": "*vcard:Contact*"

"*vcard:fn*": "*ex:Maia - Portal de dados*"

"*vcard:hasEmail*": ""

O termo "*comm*" seria substituído por "*referencedby*", que faz referência a uma publicação ou cita um recurso catalogado.

Existe também um termo designado por "*qualifiedRelation*" que apresenta um link para uma descrição de um relacionamento com outro recurso que poderá ser útil para a câmara.

No contexto mencionado, o termo "fonte" não está incluído no *DCAT-AP*, ele foi introduzido apenas numa versão recente de junho deste ano.

Assim, ao revisar esses metadados, sugere-se substituir o termo utilizado anteriormente por "accessService", que proporciona acesso à distribuição do conjunto de dados. Ele inclui o "downloadURL", um URL para baixar o conjunto de dados em um formato específico, substituindo o termo 'exemplo'. Além disso, o termo "api url" deveria ser alterado para "accessURL", que é o URL do recurso que permite acesso aos dados via web, consulta ou API. Quando o conjunto de dados estiver disponível online, pode-se usar o termo "landingPage".

Seria importante a existência de uma classe adicional designada "soma de verificação" ("Checksum"), que permitisse verificar a integridade dos dados durante o armazenamento, demonstrando assim a autenticidade do conteúdo. O termo "algoritmo" ("algorithm") identifica o algoritmo utilizado para gerar o valor.

Português	Inglês	Explicação
Título	title	Nome dado ao recurso
Descrição	description	Texto sobre o recurso
Criador	creator	Entidade responsável pela produção
Referenciado por	referencedby	Recurso relacionado/ faz referência ou cita
Data de início	startDate	Início do período
Data de Fim	endDate	Fim do período
Alteração/ Modificação de data	modified	Data mais recente que foi alterada, atualizada ou modificada
Data emissão	issued	Data emissão formal do recurso
Tema	theme	Tema que se refere
Spatial	espacial	Área geográfica coberta pelos dados
Licença	license	Documento no qual o recurso é disponibilizado
Editor	publisher	Entidade responsável por disponibilizar o recurso

Ponto de contato	contactPoint	Informações de contacto importantes para o recurso
Acesso de serviço	accessService	Serviço de dados que possibilita o acesso à distribuição dos dados
Acesso Url	accessUrl	Utilizado para serviço que disponibiliza acesso a distribuição por meio formulário web ou API
download Url	downloadURL	Links diretos para recursos
Página de destino	landingPage	Página web que pode ser navegada e oferece acesso aum conjunto de dados ou informações adicionais.

Tabela 2- Adaptação dos mais Metadados conforme DCAT-AP

Objetivos da Solução Proposta:

- Conformidade com *DCAT-AP*: Alinhar os metadados dos conjuntos de dados da Câmara Municipal com as especificações do *DCAT-AP* para garantir a interoperabilidade.
- Qualidade dos Metadados: Melhorar a qualidade dos metadados, assegurando que sejam completos, precisos e atualizados.
- Integridade dos Dados: Garantir que os dados disponibilizados sejam íntegros e confiáveis, proporcionando uma base sólida para a tomada de decisões informadas

A implementação da solução será guiada pelas seguintes diretrizes:

- Atributos Obrigatórios: Garantir que todos os atributos obrigatórios definidos pelo *DCAT-AP* estejam presentes e corretamente preenchidos. Isto inclui títulos, descrições, datas de publicação, formatos de dados, entre outros.

Melhores Práticas de Gestão de Metadados:

- Consistência e Padronização: Adotar terminologia e formatos padronizados para garantir a consistência dos metadados.
- Validação e Verificação: Implementar processos de validação e verificação para assegurar que os metadados sejam corretos e estejam em conformidade com as especificações do *DCAT-AP*.

6. Conclusões e reflexão crítica

Durante o projeto de adaptação dos metadados para conformidade com o DCAT-AP, foram implementados ajustes significativos para melhorar a interoperabilidade e qualidade dos dados catalogados.

6.1. Lições aprendidas

A realização deste projeto proporcionou diversas aprendizagens valiosas. Em primeiro lugar, a familiarização com o padrão *DCAT-AP* e sua aplicação prática demonstrou a importância de padrões internacionais para a interoperabilidade e a qualidade dos dados. O processo de adaptação dos metadados destacou a necessidade de uma abordagem metódica e detalhada para garantir a conformidade.

A aplicação prática dos conceitos estudados no curso revelou-se fundamental para o sucesso do projeto. O conhecimento prévio sobre gestão de dados, padrões de metadados e a utilização de plataformas como o *OpenDataSoft* foram cruciais.

6.2. Trabalho futuro

Para futuras etapas do projeto, recomenda-se a continuação da avaliação e ajuste dos metadados para assegurar a manutenção da conformidade com o *DCAT-AP* à medida que novos conjuntos de dados são adicionados ao catálogo. Também seria benéfico explorar ferramentas automatizadas de verificação de conformidade para agilizar o processo de validação dos metadados.

Por fim, implementar esses metadados padronizados na plataforma *OpenDataSoft*, ajustando e atualizando continuamente conforme necessário, garantirá a manutenção da qualidade e a relevância dos dados públicos disponibilizados.

Referências bibliográficas

ALVES, R. C. V. (2010). Metadados como elementos do processo de catalogação, Faculdade de Filosofia e Ciências – Universidade Estadual Paulista.

Milagre, J. and J. Segundo (2015). A propriedade dos dados e a privacidade na perspectiva da Ciência da Informação. revista eletrônica de biblioteconomia e ciência da informação. 20: 47-76.

Sant'Ana, R. C. G. (2016). CICLO DE VIDA DOS DADOS: UMA PERSPECTIVA A PARTIR DA CIÊNCIA DA INFORMAÇÃO. Londrina. 21: 116-142.

Ribeiro, T. (2019). ISO 37120 e Objetivo do Desenvolvimento Sustentável 11: Convergência frente à Agenda 2030, Universidade Nove de Julho.

Miloslavskaya, N. and A. Tolstoy (2016). "Big Data, Fast Data and Data Lake Concepts." ELSEVIER 88: 300-305.

Stefani, S., A. Procidonio, L. Raifur and C. Chiusoli (2023). CIDADES SUSTENTÁVEIS E ISO 37120: A VISÃO DOS MUNICÍPIOS. Boletim de Conjuntura 15.

Opendatasoft. (2024). "Create the best data experiences- opendatasoft." from <https://www.opendatasoft.com/en/>.

Maia, C. M. d. from <https://www.cm-maia.pt/>.

Commission, E. "DCAT Application Profile for data portals in Europe." from https://ec.europa.eu/isa2/solutions/dcat-application-profile-data-portals-europe_en/.

República, D. d. (2016). Lei n.º 26/2016, de 22 de agosto. Portugal Diário da República 2777 - 2788.

Dekkers, M., S. Kotoglou, C. Nelson, M. Pellegrino, N. Hohn and V. Peristeras "StatDCAT-AP A Common Layer for the Exchange of Statistical Metadata in Open Data Portals."

Oliveira, L. E. R. d. A. (2017). UM MODELO DE ARQUITETURA PARA SISTEMAS GERENCIADORES DE DADOS NA WEB, Universidade Federal de Pernambuco.

opendatasoft. "Introduction à l'API Explore." From <https://userguide.opendatasoft.com/fr/article/b3dyxplacz-une-introduction-l-api-explor>.

Commission, E. "StatDCAT-AP – DCAT Application Profile for description of statistical datasets Version 1.0.1." from

https://joinup.ec.europa.eu/sites/default/files/distribution/access_url/2019-05/0812e528-c428-4832-b674-d5b9c68d1b42/StatDCAT-AP_1.0.1.pdf.

Commission, E. (2016). "GeoDCAT-AP: A geospatial extension for the DCAT application profile for data portals in Europe." from https://semiceu.github.io/GeoDCAT-AP/releases/1.0.1/geodcat-ap_1.0.1.pdf.

Europeia, J. O. d. U. (2023). REGULAMENTO DE EXECUÇÃO (UE) 2023/138 DA COMISSÃO de 21 de dezembro de 2022 que estabelece uma lista de conjuntos específicos de dados de elevado valor e as disposições relativas à respetiva publicação e reutilização: 33.

Commission, E. (2023). "INSPIRAR Infraestrutura para Informação Espacial na Europa." from https://knowledge-base.inspire.ec.europa.eu/index_en.

Europeia, J. O. d. U. (2007). DIRECTIVA 2007/2/CE DO PARLAMENTO EUROPEU E DO CONSELHO de 14 de Março de 2007 que estabelece uma infra-estrutura de informação geográfica na Comunidade Europeia (Inspire).

Albertoni, R., D. Browning, S. Cox, A. G. Beltran, A. Perego and P. Winstanley. (2020). "Vocabulário de Catálogo de Dados (DCAT) - Versão 2." from <https://www.w3.org/TR/vocab-dcat/#toc>.

Platform, O. B. "ISO 19115-1:2014." from <https://www.iso.org/obp/ui/en/#iso:std:53798:en>.

NEUMAIER, S., J. UMBRICH and A. POLLERES (2016). "Automated Quality Assessment of Metadata across Open Data Portals ": 30.

EUROPEIA, C. (2023). REGULAMENTO DE EXECUÇÃO (UE) 2023/138 DA COMISSÃO de 21 de dezembro de 2022 que estabelece uma lista de conjuntos específicos de dados de elevado valor e as disposições relativas à respetiva publicação e reutilização: 33.

Ferreira, J. A. (2013). " O MODELO DE DADOS RESOURCE DESCRIPTION FRAMEWORK (RDF) E O SEU PAPEL NA DESCRIÇÃO DE RECURSOS ": 12-23.

Albertoni, R., D. Browning, S. J. Cox, A. G. Beltran, A. Perego and P. Winstanley. (2024). "Vocabulário de Catálogo de Dados (DCAT) - Versão 3." from <https://www.w3.org/TR/vocab-dcat-3/#qualified-attribution>.

Cerveira, M. E. METAINFORMAÇÃO DESCRITIVA.

Nuffelen, B. V. (2024). "DCAT-AP 3.0." from <https://semiceu.github.io/DCAT-AP/releases/3.0.0/#CatalogueRecord>

Anexos

Anexo I

O Regulamento de Execução (UE) 2023/138 da comissão de 21 de dezembro de 2022, estabelece uma lista de conjuntos de dados de elevado valor por categorias e as disposições relativas à respetiva publicação e reutilização, tais como:

Categoria de Dados	Exemplos de conjuntos de dados por Categoria
Geoespacial	Unidades administrativas, toponímia, endereços, edifícios, parcelas cadastrais, parcelas de referência e parcelas agrícolas,
Observação da Terra e do ambiente	Hidrografia, sítios protegidos, altitude, geologia, ocupação do solo, ortoimagens, zonas de gestão/ restrição, regiões biogeográficas, recursos energéticos, instalações de monitorização do ambiente, habitats de monitorização do ambiente, habitats e biótopos, uso do solo, recursos minerais, zonas de risco natural, características oceanográficas, instalações industriais e de produção, regiões marinhas, solo e distribuição das espécies.
Metrológica	Dados de observações medidos pelas estações meteorológicas, dados climáticos: observações válidas, alertas meteorológicos, dados de radar e dados do modelo de previsão meteorológica numérica.
Estatística	Produção industrial, fluxo de turísticos na Europa, índices harmonizadas de preços no consumidor, contas nacionais, despesas e receitas das administrações públicas, dívida bruta consolidada das administrações públicas, contas e estatísticas ambientais, população, fecundidade e mortalidade, despesas correntes com cuidados de saúde, pobreza e desigualdade, emprego, desemprego e mão de obra potencial.
Empresas e propriedades de empresas	Informações de base sobre as empresas: principais atributos.
	Caraterísticas do canal navegável, obstruções Dados estáticos: prolongadas no canal navegável e fiabilidade, taxas de utilização das infraestruturas da via navegável, outras limitações físicas de via navegável, horário normal

Mobilidade	<p>de serviço de eclusas e pontes, localização e características dos portos e dos terminais de transbordo, listas das ajudas à navegação e sinais de trânsito, regras e recomendações de navegação.</p> <p>Dados dinâmicos: curvas isobatimétricas do canal navegável, obstruções temporárias no canal navegável, níveis de águas atuais e previsíveis nos fluviómetros, estados dos rios, canais, eclusas e pontes, restrição por motivo de cheia ou presença de gelo, alteração a curto prazo de serviço de eclusas e pontes e alteração a curto prazo das ajudas à navegação.</p> <p>cartas náuticas eletrônicas de navegação interior: eixo fluvial, ligações aos ficheiros xml externos, com indicação dos horários de funcionamento das estruturas que condicionam o tráfego, localização dos portos e dos terminais de transbordo, dados de referência dos fluviómetros de interesse para a navegação, margem da via navegável ao nível médio da água, construção nas margens, contornos de eclusas e barragens, delimitações da via/canal navegável, perigos isolados submersos e à superfície no canal/via navegável e ajudas à navegação oficiais.</p>
------------	--

Dados Governamentais abertos

Os dados governamentais abertos (OGD) relacionam-se com outros conceitos como governo e dados abertos. A sua base principal é possibilitar que os dados oficiais sejam públicos para que se possam utilizar e reutilizar para que haja um livre acesso. Estes dados fornecem informações úteis para os cidadãos. Segundo Linders, o “seu objetivo é melhorar a responsabilização pública, a participação dos cidadãos e as parcerias colaborativas intersetoriais”(LINDERS, 2013).

No ano de 2007, em Sebastopol, Califórnia, um grupo de 30 defensores do governo aberto se reuniu para desenvolver os princípios fundamentais dos dados governamentais abertos:

- Completo: os dados devem ser abrangentes, sem restrições como privacidade, segurança ou privilégio;
- Primário: os dados são provenientes da fonte original;

- Oportuno: os dados devem ser disponibilizados de forma rápida e eficiente para preservar seu valor;
- Acessível: os dados devem ser disponíveis para todos utilizadores;
- Processável por máquina: os dados devem ser estruturados e formatados para facilitar o processamento automático;
- Não discriminatório: dados disponíveis para qualquer pessoa, sem discriminação;
- Não proprietário: nenhuma entidade tem controle exclusivo dos dados;
- Livre de licença: os dados devem ser livres de restrições de patentes, marcas registradas e direitos autorais.

Para concluir, no contexto do setor público, a transparência e a disponibilização da informação são pilares fundamentais. É essencial que a população possa aceder de forma simples e fácil aos dados públicos, possibilitando uma maior compreensão das políticas e ações governamentais.

Anexo 2

Constituição de metadados obrigatórios pelas seis classes principais do *DCAT*.

Classe:catálogo	Class:catalog
Registro de catálogo	catalog record
Recurso	resource
Conjunto de dados	dataset
Serviço	service
Catálogo	catalog
Página Inicial	homepage
Temas	themes

Classe: Recurso Catalogado	Class:resource
-----------------------------------	-----------------------

Direitos de acesso	access rights
Conforme a	conforms to
Ponto de contacto	contact point
Criador	creator
Descrição	description
Política	has policy
Identificador	identifier
Referenciado por	is referenced by
Palavra-chave	keyword/tag
Página de destino	landing page
Licença	license
Linguagem	language
Relação	relation
Direitos	rights
Relação qualificada	qualified relation
Editor	publisher
Data de lançamento	release date
Tema/categoria	theme/category
Título	title
Tipo/gênero	type/genre
Data de atualização/modificação	update/modification date
Atribuição qualificada	qualified attribution
Versão atual	has current version

Tem Versão	has version
Versão anterior	previous version
Substitui	replaces
Versão	version
Notas de versão	version notes
Durar	last
Anterior	previous

Classe: conjunto de dados	Class:dataset
Distribuição	distribution
Frequência	frequency
Cobertura espacial/geográfica	spatial/geographic coverage
Resolução espacial	spatial resolution
Cobertura temporal	temporal coverage
Resolução temporária	temporal resolution
Foi gerado por	was generated by

As propriedades dcat:resource também estão disponíveis;

Classe: Registo de catálogo	Class: Catalog Record
Conforme a	conforms to
Descrição	description
Data de listagem	listing date
Tópico principal	primary topic

Título	title
Data de atualização/modificação	update/modification date

Classe: distribuição	Class: Distribution
Direitos de acesso	access rights
URL de acesso	access URL
Serviço de acesso	access service
Tamanho de bytes	byte size
Formato de compressão	compression format
Conforme	conforms to
Descrição	description
URL de download	download URL
Formato	format
Política	has policy
Licença	license
Tipo de mídia	media type
Formato ex: arquivo TAR, ZIP	packaging format
Data de lançamento	release date
Direitos	rights
Resolução espacial	spatial resolution
Resolução temporária	temporal resolution
Título	title

Data de atualização/modificação	update/modification date
---------------------------------	--------------------------

Classe: serviço de dados	Classe: Data Service
Descrição do endpoint	endpoint description
URL do endpoint	endpoint URL
Servir conjunto de dados	serves dataset

As propriedades dcat:resource também estão disponíveis

Anexo 3

Neste anexo, são apresentados os elementos dos metadados INSPIRE e geoespaciais da ISO 19115, bem como todos os elementos do GeoDCAT-AP.

Metadados INSPIRE:

- Identificador INSPIRE: Um identificador único para o dataset no âmbito da Diretiva INSPIRE.
- Tema INSPIRE: O tema INSPIRE do dataset.
- Tipo de recurso INSPIRE: O tipo de recurso INSPIRE no dataset (por exemplo, cobertura).
- Nível de hierarquia INSPIRE: O nível de hierarquia INSPIRE do dataset (por exemplo, local, regional, nacional).
- Conformidade INSPIRE: O nível de conformidade do dataset com a Diretiva INSPIRE (por exemplo, básico, completo).

Metadados Geoespaciais da ISO 19115:

- Extensão espacial: área geográfica coberta pelo dataset.
- Sistema de referência espacial: sistema de referência espacial usado no dataset (por exemplo, WGS84 (é um sistema de coordenadas em grau de latitude e de longitude), EPSG:3857(usado para maps online, coordenadas são dadas em metros)).

Adopção da plataforma OpenDataSoft para publicação de dados abertos

- Escala: escala do dataset.
- Qualidade dos dados: qualidade dos dados no dataset
- Tipo de geometria: tipo de geometria usada no dataset (por exemplo, ponto, linha, polígono).
- Tema: tema do dataset (por exemplo, meio ambiente, transporte, agricultura)
- Tipo de recurso: tipo de recurso no dataset (por exemplo, cobertura,)
- Nível de hierarquia: nível de hierarquia do dataset (por exemplo, local, regional, nacional).

Na coluna "*Obligation*", a letra M indica elementos de metadados obrigatórios, a letra C representa elementos condicionais e a letra O indica elementos opcionais (sendo condicional quando obrigatórios sob determinadas condições).

INSPIRE	Obligation	ISO 19115 Core	Obligation	DCAT-AP	GeoDCAT-AP
Metadata point of contact	M	Metadata point of contact	M		Yes
Metadata date	M	Metadata date stamp	M	Yes	Yes
Metadata language	M	Metadata language	C	Yes	Yes
		Metadata character set	C		Yes
		Metadata file identifier	O		Yes
		Metadata standard name	O		Yes
		Metadata standard version	O		Yes
Resource title	M	Dataset title	M	Yes	Yes
Temporal reference - Date of creation / publication / last revision	C	Dataset reference date	M	Partially (creation date not included)	Yes
Resource abstract	M	Abstract describing the dataset	M	Yes	Yes

Fonte: GeoDCAT-AP: A geospatial extension for the DCAT application profile for data portals in Europe Version 1.0.1¹¹

¹¹ https://semiceu.github.io/GeoDCAT-AP/releases/1.0.1/geodcat-ap_1.0.1.pdf

Adopção da plataforma OpenDataSoft para publicação de dados abertos

INSPIRE	Obligation	ISO 19115 Core	Obligation	DCAT-AP	GeoDCAT-AP
Resource language	C	Dataset language	M	Yes	Yes
Topic category	M	Dataset topic category	M		Yes
Geographic bounding box	M	Geographic location of the dataset (by four coordinates or by geographic identifier)	C	Yes	Yes
Character encoding	C	Dataset character set	C		Yes
Temporal reference - Temporal extent	C	Additional extent information for the dataset (vertical and temporal)	O	Partially (temporal extent only)	Partially (temporal extent only)
Lineage	M	Lineage	O	Yes	Yes
Spatial representation type	M	Spatial representation type	O		Yes
Encoding	M	Distribution format	O	Yes	Yes
Spatial resolution	C	Spatial resolution of the dataset	O		Yes (but as free text)
Responsible organisation	M	Dataset responsible party	O	Partially (only 3 of the 11 responsible party roles are supported)	Yes
Resource locator	C	On-line resource	O	Yes	Yes
Coordinate reference system; Temporal reference system	M; C	Reference system	O		Yes
Conformity	M			Yes	Yes

INSPIRE	Obligation	ISO 19115 Core	Obligation	DCAT-AP	GeoDCAT-AP
Resource language	C	Dataset language	M	Yes	Yes
Topic category	M	Dataset topic category	M		Yes
Geographic bounding box	M	Geographic location of the dataset (by four coordinates or by geographic identifier)	C	Yes	Yes
Character encoding	C	Dataset character set	C		Yes
Temporal reference - Temporal extent	C	Additional extent information for the dataset (vertical and temporal)	O	Partially (temporal extent only)	Partially (temporal extent only)
Lineage	M	Lineage	O	Yes	Yes
Spatial representation type	M	Spatial representation type	O		Yes
Encoding	M	Distribution format	O	Yes	Yes
Spatial resolution	C	Spatial resolution of the dataset	O		Yes (but as free text)

Fonte: GeoDCAT-AP: A geospatial extension for the DCAT application profile for data portals in Europe Version 1.0.1

Adopção da plataforma OpenDataSoft para publicação de dados abertos

Na tabela seguinte lista os elementos de metadados, com estrela (*) são usados para indicar o elemento de metadados correspondente no perfil principal da ISO 19115:2003. Para cada elemento de metadados é indicado se o elemento é mandatory (M), optional (O), conditional (C), or recommended (R) in either specification.

INSPIRE metadata *ISO19115:2003 Core Profile	Property	Domain	Range
Resource type (M) *not in ISO19115	dct:type	- (dcat:Dataset (M), dcat:Catalog (M))	rdfs:Class (skos:Concept)

INSPIRE metadata *ISO19115:2003 Core Profile	Property	Domain	Range
Topic category (M) *Dataset topic category (M)	dct:subject	- (dcat:Dataset (M))	- (skos:Concept)
Spatial data service type (M) *not in ISO19115	dct:type	- (dcat:Dataset (M), dcat:Catalog (M), dctype:Service)	rdfs:Class (skos:Concept)
Keyword value (M) *not in ISO19115 core	dc:subject dct:subject dct:type	- (dcat:Catalog (M))	- (rdfs:Literal) - (rdfs:Resource) rdfs:Class (skos:Concept)
Originating controlled vocabulary (C) *not in ISO19115 core	skos:inScheme	skos:Concept	skos:ConceptScheme
Date of creation (C) *Dataset reference date (M) - creation	dct:created	- (dcat:Dataset (M), dcat:Catalog (M))	xsd:date
Spatial resolution (C) *Spatial resolution of the dataset (O)	rdfs:comment	-	rdfs:Literal
Conformity (M) *not in ISO19115 core	prov:wasUsedBy	prov:Entity (dcat:DataSet (M))	prov:Activity
Conformity Specification (M) *not in ISO19115 core	prov:wasDerivedFrom	prov:Entity	prov:Entity
Conformity degree (M) *not in ISO19115 core	prov:generated	prov:Activity	prov:Entity

INSPIRE metadata *ISO19115:2003 Core Profile	Property	Domain	Range
Topological Consistency (C) *not in ISO19115 core	-	-	-
Data Quality – Logical Consistency – Conceptual Consistency, Domain Consistency – (R) *not in ISO19115 core	-	-	-
Responsible party (M) *Dataset responsible party (O)	dct:creator	- (dcat:Dataset (M))	dct:Agent
	prov:qualifiedAttribution	prov:Entity (dcat:Dataset (M))	prov:Attribution
Responsible party role (M)	dct:type	- (prov:Attribution)	rdfs:Class
*Metadata file identifier (O)	dct:identifier	- (dcat:CatalogRecord (O))	rdfs:Literal
Metadata point of contact (M) *Metadata point of contact (M)	prov:qualifiedAttribution	- (dcat:CatalogRecord (O))	prov:Attribution
	dcat:contactPoint	- (dcat:CatalogRecord (O))	vcard:Kind
*Metadata character set (C)	cnt:characterEncoding	cnt:Content (dcat:CatalogRecord (O))	rdfs:Literal
Coordinate Reference System (M) *Reference System (O)	dct:conformsTo	- (dcat:Dataset)	dct:Standard

Fonte: GeoDCAT-AP: A geospatial extension for the DCAT application profile for data portals in Europe Version 1.0.1

INSPIRE metadata *ISO19115:2003 Core Profile	Property	Domain	Range
Topological Consistency (C) *not in ISO19115 core	-	-	-
Data Quality – Logical Consistency -	-	-	-
Conceptual Consistency, Domain Consistency – (R) *not in ISO19115 core	-	-	-
Responsible party (M) *Dataset responsible party (O)	dct:creator	- (dcat:Dataset (M))	dct:Agent
	prov:qualifiedAttribution	prov:Entity (dcat:Dataset (M))	prov:Attribution
Responsible party role (M)	dct:type	- (prov:Attribution)	rdfs:Class
*Metadata file identifier (O)	dct:identifier	- (dcat:CatalogRecord (O))	rdfs:Literal
Metadata point of contact (M) *Metadata point of contact (M)	prov:qualifiedAttribution	- (dcat:CatalogRecord (O))	prov:Attribution
	dcat:contactPoint	- (dcat:CatalogRecord (O))	vcard:Kind
*Metadata character set (C)	cnt:characterEncoding	cnt:Content (dcat:CatalogRecord (O))	rdfs:Literal
Coordinate Reference System (M) *Reference System (O)	dct:conformsTo	- (dcat:Dataset	dct:Standard

Exemplo
<p># Metadados de recursos no GeoDCAT-AP usando uma caixa delimitadora geográfica. Este exemplo usa múltiplas codificações, ou seja, o</p> <p># recomendados (WKT e GML), além de GeoJSON. Para denotar o tipo de dados do literal GeoJSON, o URL do tipo de mídia # IANA correspondente.</p> <pre> [] dct:especial [a dct:Local ; locn:geometria "POLIGONO((-10,58 70,09,34,59 70,09,34,59 34,56,-10,58 34,56, -10,58 70,09))"^^gsp:wktLiteral ; locn:geometry "<gml:Envelope srsName='http:// www.opengis.net/def/crs/OGC/1.3/CRS84'"> <gml:lowerCorner>34,56 -10,58</gml:lowerCorner> <gml:upperCorner>70,09 34,59</gml:upperCorner> </gml:Envelope>"^^gsp:gmlLiteral] ; locn:geometry ("{"type":"Poligono","crs":{"type":"nome","propriedades":{"nome":"você n:ogc:def:crs:OGC:1.3:CRS84"}}, "coordenadas":[[[- 10,58,70,09],[34,59,70,09],[34,59,34,56],[-10,58,34,56],[-10,58,70,09]]]""^<https://www.iana.org/assignments/media-types/application/ vnd.geo+json>]. </pre> <p># Metadados de recursos no GeoDCAT-AP usando um identificador geográfico</p> <p>#Se um URI for usado para o identificador geográfico (recomendado)</p> <pre> [] </pre>

Fonte:GeoDCAT-AP: A geospatial extension for the DCAT application profile for data portals
in Europe Version 1.0.1

Anexo 4

Neste anexo, são ilustradas as propriedades do StatDCAT-AP

Novas propriedades no namespace StatDACT-AP:

Property	Attribute
URI	stat:attribute
Range	qb:AttributeProperty, expressed as a URI.
Definition	A component used to qualify and interpret observed values
Comment	Attributes enable specification of the units of measure, any scaling factors and metadata such as the status of the observation (e.g. estimated, provisional).

Property	Dimension
URI	stat:dimension
Range	qb:DimensionProperty, expressed as a URI.
Definition	A component that identifies observations
Comment	Examples of dimensions include the time to which the observation refers, or a geographic region which the observation covers.

Fonte: StatDCAT-AP – DCAT Application Profile for description of statistical datasets
Version 1.0.1

Foi criado um novo requisito para o número de séries de dados que pode conter três regiões com três valores para cada. Já, para ser visto a unidade de medida foi apresentado um novo requisito no namespace StatDCAT-AP. Uma série temporal, é descrita para ser compatível com StatDCAT-AP como o exemplo abaixo :

dct:accrualPeriodicity

<http://publications.europa.eu/resource/authority/frequency/QUARTERLY> ;

dct:temporal [schema:startDate "2006-07-01"^^xsd:date ;

schema:endDate "2016-06-30"^^xsd:date]

O StatDCAT-AP adiciona propriedades para os conjuntos de dados.

No **catálogo** as propriedades obrigatórias são:

Property	URI	Range	Usage note
dataset	dcat:dataset	dcat:Dataset	This property links the Catalogue with a Dataset that is part of the Catalogue.
description	dct:description	rdfs:Literal	This property contains a free-text account of the Catalogue. This property can be repeated for parallel language versions of the description. For further information on multilingual issues, please refer to section 0.
publisher	dct:publisher	foaf:Agent	This property refers to an entity (organisation) responsible for making the Catalogue available.
title	dct:title	rdfs:Literal	This property contains a name given to the Catalogue. This property can be repeated for parallel language versions of the name.

Catálogo: Propriedades recomendadas:

Property	URI	Range	Usage note
homepage	foaf:homepage	foaf:Document	This property refers to a web page that acts as the main page for the Catalogue.
language	dct:language	dct:LinguisticSystem	This property refers to a language used in the textual metadata describing titles, descriptions, etc. of the Datasets in the Catalogue. This property can be repeated if the metadata is provided in multiple languages.
licence	dct:license	dct:LicenseDocument	This property refers to the licence under which the Catalogue can be used or reused.
release date	dct:issued	rdfs:Literal typed as xsd:date or xsd:dateTime	This property contains the date of formal issuance (e.g., publication) of the Catalogue.
themes	dcat:themeTaxonomy	skos:ConceptScheme	This property refers to a knowledge organisation system used to classify the Catalogue's Datasets.
update/ modification date	dct:modified	rdfs:Literal typed as xsd:date or xsd:dateTime	This property contains the most recent date on which the Catalogue was modified.

Registo do Catálogo: Propriedades obrigatórias:

Property	URI	Range	Usage note
primary topic	foaf:primaryTopic	dcat:Dataset	This property links the Catalogue Record to the Dataset described in the record.
update/ modification date	dct:modified	rdfs:Literal typed as xsd:date or xsd:dateTime	This property contains the most recent date on which the Catalogue entry was changed or modified.

Registo do Catálogo Propriedades recomendadas:

Property	URI	Range	Usage note
application profile	dct:conformsTo	rdfs:Resource	This property refers to an Application Profile that the Dataset's metadata conforms to.
change type	adms:status	skos:Concept	This property refers to the type of the <i>latest</i> revision of a Dataset's entry in the Catalogue. It MUST take one of the values :created, :updated or :deleted depending on whether this <i>latest</i> revision is a result of a creation, update or deletion.
listing date	dct:issued	rdfs:Literal typed as xsd:date or xsd:dateTime	This property contains the date on which the description of the Dataset was included in the Catalogue.

Conjunto de Dados: Propriedades obrigatórias:

Property	URI	Range	Usage note
description	dct:description	rdfs:Literal	This property contains a free-text account of the Dataset. This property can be repeated for parallel language versions of the description.
title	dct:title	rdfs:Literal	This property contains a name given to the Dataset. This property can be repeated for parallel language versions of the name.

Conjunto de Dados: Propriedades recomendadas :

Property	URI	Range	Usage note
contact point	dcat:contactPoint	vcard:Kind	This property contains contact information that can be used for sending comments about the Dataset.
dataset distribution	dcat:distribution	dcat:Distribution	This property links the Dataset to an available Distribution.

Property	URI	Range	Usage note
keyword/ tag	dcat:keyword	rdfs:Literal	This property contains a keyword or tag describing the Dataset.
publisher	dct:publisher	foaf:Agent	This property refers to an entity (organisation) responsible for making the Dataset available.
theme/ category	dcat:theme, subproperty of dct:subject	skos:Concept	This property refers to a category of the Dataset. A Dataset may be associated with multiple themes.

Fonte: StatDCAT-AP – DCAT Application Profile for description of statistical datasets Version

Anexo 5

Neste anexo, são apresentadas tarefas realizadas várias tarefas na plataforma *OpenDataSoft*, incluindo a inserção dos links do Power BI nos datasets, a verificação ou correção dos títulos, a verificação dos temas atribuídos, a ordenação das tabelas por ordem crescente, desativar visualização Standard: e a verificação dos temas *ODS ICONS*.

DATASETS > DIPLOMADOS-NO-NO-ENSINO-SUPERIOR-POR-100000-HABITANTES-COPY

Consumo Final de Energia per Ca

DATASETS > DIPLOMADOS-NO-NO-ENSINO-SUPERIOR-POR-100000-HABITANTES-COPY

Consumo Final de Energia per Ca

Save Pu

Sources Processing Information **Visualizations** Export History Scheduling Security Lineage

Table Map Analyze Images Calendar Custom view

You can customize the displayed fields in the table tab, without altering your dataset.

Default sort: t (integer) Ascending order

Displayed fields: t v

t	v
2,011	36.64
2,012	35.51
2,013	34.66
2,014	34.93
2,015	35.01
2,016	36.19
2,017	36.3
2,018	36.87
2,019	36.3
2,020	35.93
2,021	34.65
2,022	
2,023	
2,024	

Sources Processing Information **Visualizations** Export History Scheduling Security Lineage

Table **Analyze** Images Calendar Custom view

☐ Enable analyze view

Adopção da plataforma OpenDataSoft para publicação de dados abertos

DATASETS > DIPLOMADOS-NO-NO-ENSINO-SUPRIOR-POR-100000-HABITANTES-COPY

Consumo Final de Energia per Ca

Save

Sources Processing **Information** Visualizations Export History Scheduling Security Lineage

Standard DCAT-AP D4M Attachments

General information

Technical identifier

diplomados-no-no-ensino-superior-por-100000-habitantes-copy Edit

Description DCAT-AP

</> 🔍 **B** *I* S ≡ ☰ ☷ 🖼️ 🔗 —

Themes

Energia ✕

Keywords DCAT-AP

Start typing to get suggestions

DATASETS **NÚMERO DE ENFERMEIROS POR 10000 HABITANTES**

Número de Enfermeiros por 100.

PUBLISHED 3 DAYS AGO MODIFIED 3 DAYS AGO

[Save](#) [Publish](#) [Explore](#) [More Actions ▾](#)

- Undo
- Redo
- Duplicate dataset**
- Unpublish
- Delete dataset

Sources Processing Information Visualizations Export History Scheduling Security Lineage

Add a new source

HTTP link
<http://baze2.cm-maia.pt/D4CMMaia/api4sV3.php?nome=NumEnfCMilHabit>
 Modified May 2, 2024 JSON File

main

Go to portal

Catalog

Datasets

Harvesters

Connections

Metadata templates

Page editor

Assets

Reuses

Look & Feel

Users & Groups

Analytics

Workspaces

Configuration

Plans

PT

Terms and conditions

Release notes

DATASETS > NÚMERO-DE-ENFERMEIROS-POR-100000-HABITANTES-COPY

Número de Enfermeiros por 100.

Save Publish Explore More Actions

Sources Processing Information Visualizations Export History Scheduling Security Lineage

Standard DCAT-AP D4M Attachments

General information

Technical identifier

numero-de-enfermeiros-por-100000-habitantes-copy Edit

Description DCAT-AP

</> Bold Italic Link Unlink Bulleted List Numbered List Table Media Embed Remove

Número de Enfermeiros por 100.000 Habitantes

Themes

Saúde

Keywords DCAT-AP

Start typing to get suggestions

Anexo 6

Este anexo descreve uma análise realizada antes da adequação dos metadados. Assim, é descrita a importância de cada campo que compõe o catálogo inicial da Câmara Municipal da Maia.

- Sigla/nome- a sigla do assunto, facilita a identificação rápida e única.
- Indicador- o nome específico do indicador define de forma clara o que a sigla pretende representar.
- Descrição- apresenta o que o indicador mede e o seu propósito, fornece o contexto.
- Tema- área a que o indicador pertence, facilita a organização e a pesquisa dentro de categorias.
- Forma de cálculo- fórmula utilizada para o cálculo.
- Fonte- origem de dados, demonstra credibilidade e a veracidade das informações. Detalha como o indicador é calculado, permitindo que os usuários reproduzam os resultados.
- Origem- detalhes mais específicos sobre como e onde os dados foram reutilizados.
 - Divisão territoriais- áreas geográficas, importante para análises comparativas (ex: freguesia, concelho, NUTS)
 - Unidade- Medida de análise utilizada (ex: número, %, m2, m3, km, km2, Ha, km3, metro, euro, kWh, °C, db(A), Ton, ug/m3, minutos, mm:ss)
 - Forma de atualização- método utilizado para atualização de dados (ex: manual, REST API, acesso de leitura DB, ficheiro manual e ficheiro automático)
 - Periodicidade de Atualização- frequência com que os dados são atualizados (ex: bienal, anual, semestral, trimestral, mensal, diário, ad-hoc)
- Observações- comentários adicionais, importantes para informações extra.
- Editor- O nome do responsável pela entrada dos dados, ajuda a garantir que as pessoas sejam responsáveis pela precisão e integridade das informações que adicionam.

NUTS- A Nomenclatura Unidade Territórios para Fins Estatísticos (NUTS) constitui um sistema hierárquico de divisão territorial. A definição NUTS foi estabelecida inicialmente pelo

Decreto-Lei nº46/89. Porém, sofreu alterações até a publicação do Decreto-Lei nº 46/89, de 15 de fevereiro, que estabelece as matrizes de delimitação das NUTS.

A NUTS é utilizada para fins estatísticos e serve de base à definição de indicadores regionais. Está subdividida em três níveis regionais distintos em cada país-membro da União Europeia (UE). Em 2024, houve alterações nas NUTS conforme o quadro abaixo:

NUTS I	NUTS II	NUTS III
Portugal Continente	Norte	Alto Minho Cávado Ave Área Metropolitana do Porto Alto Tâmega e Sousa Tâmega e Sousa Douro Terras de Trás-os-Montes
	Centro	Região de Aveiro Região de Coimbra Região de Leiria seu Dão Lafões Beira Baixa Beiras e Serra da Estrela
	Grande Lisboa	Grande Lisboa
	Oeste e Vale do Tejo	Médio Tejo Lezíria do Tejo Oeste

	Península de Setúbal	Península de Setúbal
	Alentejo	Alentejo Litoral Baixo Alentejo Alto Alentejo Alentejo Central
	Algarve	Algarve
Região Autónoma dos Açores	Região Autónoma dos Açores	Região Autónoma dos Açores
Região Autónoma da Madeira	Região Autónoma da Madeira	Região Autónoma da Madeira