

# Cadernos BAD

---

## Plataforma CLAV: contributo para a disponibilização de dados abertos da Administração Pública em Portugal<sup>1</sup>

Alexandra Lourenço

Direção-Geral do Livro, dos Arquivos e das Bibliotecas (DGLAB)

[alexandra.lourenco@dglab.gov.pt](mailto:alexandra.lourenco@dglab.gov.pt)

José Carlos Ramalho

Universidade do Minho

[jcr@di.uminho.pt](mailto:jcr@di.uminho.pt)

Maria Rita Gago

Direção-Geral do Livro, dos Arquivos e das Bibliotecas (DGLAB)

[m-rita.gago@dglab.gov.pt](mailto:m-rita.gago@dglab.gov.pt)

Pedro Penteado

Direção-Geral do Livro, dos Arquivos e das Bibliotecas (DGLAB)

[pedro.penteado@dglab.gov.pt](mailto:pedro.penteado@dglab.gov.pt)

### Resumo

A comunicação começa por indicar as políticas e estratégias, na Europa e em Portugal, para a disponibilização de dados abertos, nos domínios da ciência aberta e da Administração Pública (AP), bem como as políticas nacionais para a transformação digital neste setor. Salienta a estratégia que tem vindo a ser desenvolvida pelo organismo de coordenação da política arquivística, a Direção-Geral do Livro, dos Arquivos e das Bibliotecas (DGLAB), para a promoção da interoperabilidade semântica e de gestão da informação arquivística, nomeadamente ao nível da sua classificação e da avaliação.

Apresenta um conjunto de iniciativas da DGLAB, de que se destaca o desenvolvimento da Lista Consolidada para a classificação e avaliação da informação pública (LC), que serve de referencial para a construção normalizada dos planos de classificação e tabelas de seleção das entidades que executam funções de Estado. Especifica o projeto da DGLAB para operacionalizar a utilização da LC, intitulado “M51-CLAV – Arquivo

---

<sup>1</sup> Artigo escrito no 1.º semestre de 2019 e apresentado no IX Encontro EDICIC 2019, realizado em Barcelona. Posteriormente foi revisto e atualizado para ser proposto para publicação nos Cadernos BAD, ao abrigo do acordo estabelecido entre a BAD e a Comissão de Organização do EDICIC.

digital: Plataforma modular de classificação e avaliação da informação pública” (CLAV), enquadrado pela Medida 51 do Programa de simplificação administrativa Simplex +, intitulado “Arquivo digital”; Este Projeto permitiu concretizar a Plataforma CLAV, que disponibiliza a referida Lista Consolidada, com as funções e processos de negócio das entidades que exercem funções públicas, bem como catálogos destas entidades e da legislação associada a esses processos. Os seus dados servem para a elaboração de planos de classificação e tabelas de seleção das entidades, devendo ser integrados nos seus sistemas de informação. São ainda utilizados para a troca de informação entre entidades, promovendo a interoperabilidade semântica na AP. A Plataforma viabiliza também a desmaterialização dos procedimentos associados à elaboração das tabelas de seleção e ao controlo da eliminação da informação pública, contribuindo assim para uma maior transparência da atuação da Administração e para uma cidadania ativa.

A comunicação dá ainda conta das fases para a concretização do referido Projeto, bem como dos resultados obtidos, dando especial destaque ao modelo ontológico, fundamental para a obtenção de dados abertos, e para a sua conexão com a web semântica.

Os dados da CLAV serão disponibilizados em CSV, RDF OWL, SKOS e XML, através de uma API de dados (“Application Programming Interface”), na plataforma Dados.gov.pt e no Portal ePortugal e referenciados no catálogo europeu de ativos semânticos da Joinup (plataforma colaborativa da Comissão Europeia, inserida no Programa ISA2).

A comunicação especifica ainda a importância dos dados disponibilizados através da CLAV para a aplicação do Regulamento Geral de Proteção de Dados.

**Palavras-chave:** Administração Pública, Arquivos, Avaliação suprainstitucional, Classificação funcional, Dados abertos, Lista Consolidada, Ontologia, Plataforma CLAV

## **CLAV Platform: a contribution to provide open data for Public Administration in Portugal**

### **Abstract**

CLAV Platform: a contribution to provide access to Portuguese Public Administration open data  
Keywords: Public Administration, archives, macro appraisal, functional classification, open data, ontology, CLAV Platform  
This presentation aims to specify CLAV’s requirements, its modeling process and to exploit its data at an Open Administration context; it also aims to present the platform’s main features and to reveal how it will promote public administration open data access.

The presentation begins with an introduction to European and Portuguese strategies and policies to provide access to open data, regarding open science and public administration, as well as the current digital transformation policies. Such policies tend to promote semantic interoperability and archival information management, with special focus on classification and appraisal, supported by a new legal framework. Among other initiatives, the Portuguese coordination agency for the archival policy has developed a classification and appraisal framework, based on a public administration functional classification and business process list.

CLAV is an ontology based electronic platform, developed to implement the use of that business process list, in order to support the creation of business activity structure classification schemes and retention schedules to integrate the organizational information management systems.

The data hosted and managed by CLAV is meant to reach a diverse public and will be available in CSV, RDF, OWL, SKOS and XML, through a data API (Application Programming Interface) at Portuguese electronic data dissemination and exchange platforms such as Dados.gov and the e Portugal Portal, the former Portal do Cidadão. CLAV will also be referenced at the European Commission Joinup catalog, a collaborative platform within the ISA2 Programme – Interoperability solutions for public administrations, businesses and citizens.

The presentation specifies also the importance of the data made available through CLAV to the implementation of General Data Protection Regulation.

**Keywords:** Public Administration, archives, macro appraisal, functional classification, open data, ontology, CLAV Platform

## Introdução

Nos últimos tempos temos assistido à definição de políticas e estratégias, na Europa e em Portugal, para a disponibilização de dados abertos, nos domínios da ciência aberta e da Administração Pública (AP) (Silva, 2017). Neste último caso, a política europeia de livre acesso aos dados foi preconizada pela Diretiva 2013/37/UE do Parlamento Europeu e do Conselho, de 26 de junho de 2013, transposta para a Lei n.º 26/2016, que aprova o regime de acesso à informação administrativa e promove a sua reutilização através do uso de formatos de dados abertos. Também a Declaração de Talin, de 2017, que segue o “Plano de ação europeu (2016–2020) para a administração pública em linha”, reconheceu a necessidade de serviços públicos digitais interoperáveis e de uma maior abertura e transparência no acesso dos cidadãos aos dados recolhidos pela AP.

A nível internacional é ainda de destacar a Parceria para o Governo Aberto (Open Government Partnership), destinada a promover a transparência e a capacitar os cidadãos para uma melhor governança, a qual Portugal integrou em 2017. Neste enquadramento subscreveu a Declaração de Administração Aberta, promoveu a criação de uma Rede Nacional de Administração Aberta e apresentou um Plano de ação de Administração aberta que contemplava, entre outros, a reformulação do Portal Nacional de Dados Abertos da Administração Pública (Dados.gov), ocorrida em 2018.

Em simultâneo, em Portugal, têm sido promovidas políticas para a transformação digital na Administração Pública, com capacidade para alavancar a disponibilização de dados abertos. Foram consignadas em diplomas e documentos orientadores, como a Resolução do Conselho de Ministros (RCM) n.º 108/2017, de 26 de julho, que aprovou a Estratégia TIC 2020, com 3 eixos: 1/ integração e interoperabilidade; 2/ inovação e competitividade e 3/ partilha

de recursos e a RCM n.º 2/2018, que procedeu à revisão do Regulamento Nacional de Interoperabilidade Digital. Merece ainda destaque a RCM n.º 51/2017, de 19 de abril, que pretende reduzir o consumo de papel na Administração Pública, “promovendo a otimização de processos e a modernização de procedimentos administrativos”, nomeadamente através da desmaterialização de processos, da promoção da adoção de sistemas de gestão documental eletrónica ou outros e da digitalização de documentos destinados a ser arquivados. Entre as medidas previstas estão a adoção de processos de “classificação, avaliação e seleção de informação, tendo em consideração, sempre que possível, os princípios de uma Macroestrutura Funcional (MEF) e a Avaliação Suprainstitucional na Administração (ASIA)”. Viria a ser reajustada com a RCM n.º 141/2018 de 26 de outubro.

Neste contexto, o organismo de coordenação da política arquivística – a Direção-Geral do Livro, dos Arquivos e das Bibliotecas (DGLAB) – tem vindo a desenvolver estratégias de promoção da interoperabilidade semântica e de gestão da informação arquivística, nomeadamente ao nível do seu registo, classificação e avaliação. Para responder às necessidades neste domínio, a DGLAB levou a efeito um conjunto de iniciativas, de que se destaca a divulgação do esquema Metainformação para a Interoperabilidade (MIP), utilizado no Modelo de Dados Canónico da Plataforma de interoperabilidade na Administração Pública (iAP), bem como a publicação da Lista Consolidada para a classificação e avaliação da informação pública (LC), que serve de referencial para a construção normalizada dos planos de classificação e tabelas de seleção das entidades organizacionais que executam funções de Estado.

A Lista Consolidada é o resultado de projetos colaborativos e transversais que envolveram representantes de centenas de entidades da AP e que ocorreram sobretudo na última década. Visaram a construção de uma linguagem comum para a classificação da informação arquivística, numa perspetiva funcional, passível de utilizar no registo e na troca de documentos via iAP, bem como para a avaliação dessa informação, numa perspetiva suprainstitucional. Referimo-nos ao projeto Macroestrutura Funcional (MEF), que teve por objetivo a identificação das funções e subfunções executadas pela Administração Pública, ao projeto “Harmonização de terceiros níveis em planos de classificação conformes à MEF”, que teve por finalidade a identificação de processos de negócio em que se decompõem as funções, e ao projeto “Avaliação suprainstitucional da informação arquivística” (ASIA), que teve por objetivo a harmonização, entre as entidades públicas, dos prazos de conservação e destinos finais da documentação produzida no âmbito dos processos de negócio.

Para operacionalizar a utilização da Lista Consolidada, a DGLAB, enquadrada pela Medida 51 do Programa de simplificação administrativa Simplex +, intitulada “Arquivo digital”, procedeu, no âmbito do Aviso SAMA n.º 02/SAMA2020/2016, ao desenvolvimento do Projeto “M51 – CLAV – Arquivo digital: Plataforma modular de classificação e avaliação da informação pública” (CLAV). Para o efeito, obteve a colaboração da Universidade do Minho para os

trabalhos da componente tecnológica e da empresa Morais Leitão, Galvão Teles, Soares da Silva & Associados (MLGTS), para a elaboração da proposta de Regime jurídico da classificação e avaliação da informação arquivística (RJCAIA), de suporte à utilização da Plataforma CLAV e às mudanças de procedimentos de gestão da informação a operar na AP. A DGLAB viria ainda a elaborar uma proposta de Regulamento Geral para complementar o RJCAIA.

A Plataforma CLAV disponibiliza um conjunto de produtos e serviços orientados para a Administração Pública, as empresas e o cidadão, de que destacamos uma ontologia com os processos de negócio das entidades que exercem funções públicas (Lista Consolidada, LC), associados a um catálogo da legislação que regula esses processos e a um catálogo dos organismos que os executam e que são responsáveis pela preservação da informação produzida nesse contexto e pela gestão do seu ciclo de vida.

Integra ainda tabelas de seleção, derivadas da LC, com identificação das classes e do ciclo de vida da informação associada aos processos de negócios (prazos de conservação administrativa e destino final) para implementação em contexto organizacional ou pluriorganizacional. Os dados das tabelas de seleção, tal como sucede com os da LC, são disponibilizados em formato aberto, para facilitar a sua integração nos sistemas de informação organizacionais. São utilizados no registo e na classificação da informação, enquanto metainformação dos documentos integrados nestes sistemas, bem como para a troca destes documentos entre entidades organizacionais, promovendo a interoperabilidade semântica, por via do uso de uma linguagem comum.

Por outro lado, a Plataforma viabiliza ainda a desmaterialização dos procedimentos associados à atualização da Lista Consolidada, à elaboração das tabelas de seleção e ao controlo da eliminação da informação arquivística, através da submissão de autos diretamente na Plataforma. Esta submissão é obrigatória para todas as entidades públicas que pretendam eliminar documentos ou informação produzida no âmbito das suas atividades, de acordo com o Decreto-Lei n.º 447/88, de 10 de dezembro, e as portarias de gestão de documentos subsequentes. A Plataforma será ainda utilizada na elaboração, aprovação, atualização e controlo de aplicação de Relatórios de avaliação de documentação acumulada (RADA) da AP, pela DGLAB. Os trabalhos para este efeito só serão desenvolvidos a partir do final de 2019, pelo que não serão analisados neste artigo.

Está também previsto que estes dados possam alcançar outros públicos, através da disponibilização em formatos abertos na Plataforma Dados.gov e no Portal ePortugal – <https://eportugal.gov.pt/> (anterior Portal do cidadão), bem como a sua referência no catálogo europeu de ativos semânticos da Joinup (plataforma colaborativa da Comissão Europeia, inserida no Programa ISA2 – Interoperability solutions for public administrations, businesses and citizens).

Neste contexto, a presente comunicação tem como objetivos:

- 1) Especificar os requisitos e a modelação que permitiram a construção da Plataforma CLAV e a exploração dos seus dados em contexto de Administração aberta;
- 2) Apresentar as principais funcionalidades da Plataforma;
- 3) Identificar as formas previstas para a disponibilização de dados abertos da AP a partir da CLAV.

## Método

O método utilizado na nossa abordagem centra-se no estudo de caso da Plataforma “CLAV – Classificação e avaliação da informação pública” e na sua aplicação de princípios teórico-metodológicos sobretudo da Ciência da Computação e da Ciência da Informação.

A CLAV foi desenvolvida com base nas seguintes etapas:

- 1) **Levantamento de requisitos funcionais e não funcionais** – Esta etapa permitiu:
  - 1/ Identificar os requisitos da ontologia acima referida, de que se destaca a definição das entidades do modelo ontológico, bem como dos aspetos necessários à sua gestão;
  - 2/ Identificar os requisitos necessários à desmaterialização do procedimento de avaliação e de eliminação, incluindo procedimentos de submissão de novas propostas;
  - 3/ Identificar enquadramentos legais, serviços e produtos com os quais a CLAV fará interface como fornecedor de serviços, como por exemplo: Regulamento Nacional de Interoperabilidade Digital; Plataforma de interoperabilidade na Administração Pública (iAP), Portal Dados.gov, Linked Open Data (LOD);
  - 4/ Identificar e caracterizar os mecanismos de interação entre as entidades públicas e o cidadão para disponibilização de informação pública (em articulação com o Portal ePortugal e Portal Dados.gov);
  - 5/ Caracterizar os mecanismos de autenticação previstos nas políticas e programas de modernização administrativa.
- 2) **Modelação e análise de informação** – Esta etapa permitiu:
  - 1/ Definir o modelo abstrato de dados do sistema de informação de suporte à plataforma modular. Partiu-se da análise dos requisitos levantados na atividade anterior, para efetuar a especificação formal do modelo com todas as entidades, atributos e relações necessárias;
  - 2/ Produzir a especificação das interfaces com sistemas externos: iAP, Dados.gov, Linked Open Data;
  - 3/ Decidir sobre as tecnologias e normas usadas na implementação do sistema.
- 3) **Implementação de sistema de suporte à ontologia que comporta a Lista Consolidada** – Esta etapa permitiu desenvolver o sistema de modo a apresentar as seguintes funcionalidades para a gestão da informação ativa:
  - 1/ Disponibilização da Lista Consolidada numa forma ontológica, em formatos legíveis por humanos, HTML e PDF, e por máquinas, JSON, XML, RDF (Resource Description

Format), ou em linguagens para representação de conhecimento como SKOS (Simple Knowledge Organization System) e OWL (Ontology Web Language), com os respetivos procedimentos de autenticação e de alerta para atualização;

2/ Interface a serviços externos: iAP, Dados.gov e Linked Open Data;

3/ Interface genérica para máquinas: API (Application Programming Interface) REST (Representational State Transfer);

4/ Interface Web para humanos;

5/ Operações de manutenção da Lista Consolidada em ambiente Web colaborativo, com a respetiva autenticação e controlo de acessos e perfis de utilização;

6/ Operações de manutenção da lista de utilizadores e perfis de utilizadores;

7/ Autenticação;

8/ Operações de manutenção das listas de entidades organizacionais e de legislação associada à Lista Consolidada e tabelas de seleção específicas;

9/ Submissão de pedidos de alterações na Lista Consolidada;

10/ Disponibilização de um eixo temporal alocado à Lista Consolidada sobre o qual se possam colocar interrogações sobre:

- prazos de conservação;
- destinos finais;
- entidades organizacionais (donos e participantes dos processos de negócio);
- histórico de processos de negócio;
- histórico legislativo;
- outros campos da Lista Consolidada.

11/ Importação de informação proveniente de outros sistemas, por ex., através de ficheiro CSV.

**4) Implementação de sistema de criação de tabelas de seleção e de controlo da eliminação da informação arquivística ativa** – Esta etapa permitiu:

1/ Construção assistida e armazenamento de tabelas de seleção específicas;

2/ Disponibilização de funcionalidades que possibilitem o controlo proativo da eliminação da informação arquivística de acordo com as tabelas de seleção;

3/ Desmaterialização do atual procedimento de receção e análise de autos de eliminação remetidos pelos organismos da Administração Pública para a DGLAB, através de aplicação com interface Web que permita aos vários organismos da AP a criação assistida de autos de eliminação, ou a receção de autos preenchidos em modelo pré-definido ou enviados pelo sistema de informação da entidade através de API.

4/ Produção de indicadores estatísticos sobre os autos de eliminação.

**5) Testes, avaliação de usabilidade e implementação de melhorias** – Esta etapa dividiu-se em duas partes:

1/ Análise da migração dos dados da Lista Consolidada disponibilizada em folha de cálculo para a Plataforma;

2/ Teste aos produtos e funcionalidades da Plataforma, na qual se procedeu à verificação do cumprimento dos vários requisitos identificados, necessários ao seu funcionamento, mas também à melhoria de questões relacionadas com a usabilidade, além da implementação de todos os fluxos de trabalho, internos e externos, da autenticação de utilizadores, bem como à submissão, análise e aprovação de pedidos.

Estas etapas tiveram em conta a disponibilização de conteúdos em formatos abertos, de que adiante se dará informação mais detalhada.

## Resultados

A partir das etapas identificadas no ponto anterior, obtiveram-se os seguintes resultados:

### 1) Requisitos para as entidades do modelo ontológico e respetiva gestão

Na fase inicial do levantamento de requisitos, foram identificadas as principais entidades (classes abstratas) do modelo ontológico: Classes (Lista Consolidada), Tabela de Seleção, Auto de Eliminação, Legislação, Entidades organizacionais, Tipologia de Entidades, bem como outras de gestão (Utilizador e Logs).

Foi necessário identificar os elementos informativos destas entidades, para mais tarde os transformar em atributos ou relações entre classes abstratas.

Classes (Lista Consolidada): Estrutura hierárquica de classes criada para a classificação dos processos de negócio (PN) da AP, constituída por 4 níveis, com os seguintes atributos:

1.º nível: representação das funções da AP;

2.º nível: representação das subfunções da AP;

3.º nível: representação dos processos de negócio (PN) executados pela AP;

4.º nível: representação de subdivisão dos processos de negócio para efeitos de avaliação. Esta classe apenas foi definida nos casos em se revelou necessário atribuir diferentes prazos e destinos finais à materialização informacional de distintas etapas dos PN.

Estruturou-se ainda um interface para a disponibilização específica dos Termos de Índice associados a cada classe de 3º nível ou 4º nível, quando aplicável.

Tabela de Seleção: Gerada a partir da Lista Consolidada, representa um subconjunto da LC em que aos atributos de cada nodo são acrescentados novos atributos que permitam guardar informação relativa a prazos de conservação administrativa, forma de contagem desses prazos e destino final a aplicar à documentação/informação produzida ou recebida por entidade(s) específica(s).

Auto de Eliminação: Registo comprovativo do abate documental no contexto da aplicação da tabela de seleção, segundo as regras estabelecidas em regulamento.



Plataforma CLAV: contributo para a disponibilização de dados abertos  
da Administração Pública em Portugal

Responsabiliza a entidade organizacional proponente. É composto pelos seguintes elementos de informação:

| Zonas e Elementos de informação                   | Nota explicativa   |
|---|--|
| <b>1. Zona de autenticação</b>                    | Zona em que se identifica e procede à legitimação do auto de eliminação  |
| 1.1. N.º do auto de eliminação                    | Número de identificação do auto de eliminação. Número atribuído automaticamente no momento da submissão na Plataforma CLAV e que é composto por um número sequencial seguido do ano.   |
| 1.2. Data do auto de eliminação*                  | Data em que o auto de eliminação é submetido na Plataforma CLAV.   |
| 1.3. Entidade produtora do auto de eliminação*    | Identificação da entidade produtora do auto de eliminação. Entidade responsável pela eliminação. Pode corresponder à entidade produtora ou à que lhe sucedeu;  |
| 1.4. Identificação dos responsáveis               | Identificação do(s) responsável(eis) pela ação de eliminação e validação do auto de eliminação.<br><br>Pessoa responsável pela criação do auto. Corresponde ao utilizador que estiver autenticado no sistema para este efeito;                       |
| 1.5. Diploma que legitima a eliminação*           | [eAutorizadoPor] Relação entre o AE e o diploma legal, que deverá estar registado no catálogo legislativo da Plataforma, e que autoriza a ação. Pode ser uma PGD ou um despacho, no caso do RADA - Relatório de Avaliação de Documentação Acumulada; |
| <b>2. Zona de identificação e controlo global</b> | Zona em que se identifica de forma genérica a informação/ documentação a eliminar  |
| 2.1. Designação do Fundo                          | Identificação do conjunto documental Fundo.  |
| 2.2. Número total de agregações*                  | Número de agregações identificadas no auto de eliminação. Elemento de informação que resulta da soma das agregações identificadas em cada classe da <i>Zona de identificação e controlo por classe</i> .   |

| Zonas e Elementos de informação               | Nota explicativa   |
|---|--|
| 2.3. Dimensão total por tipo de suporte*      | <p>Dimensão dos vários tipos de suporte identificados no auto de eliminação. Elemento de informação que resulta do somatório das dimensões por tipo de suporte identificadas em cada classe da <i>Zona de identificação e controlo por classe</i>.</p> <p>A dimensão da documentação em suporte papel deve ser apresentada em metros lineares e a dimensão em suporte digital em GB.</p> |
| 3.Zona de identificação e controlo por classe | <p>Zona em que se individualiza por classe a informação / documentação a eliminar. Os elementos informativos desta zona são desdobrados por cada classe, podendo-se reportar várias classes num mesmo auto.</p>  |
| 3.1. Código da classe                         | Código da classe da informação / documentação a eliminar.  |
| 3.2. Título da classe*                        | Título da classe da informação / documentação a eliminar.  |
| 3.3. Prazo de conservação administrativa*     | Prazo de conservação administrativa da classe.   |
| 3.4. Forma de contagem do PCA*                | Forma de contagem do prazo de conservação administrativa indicada na Tabela de seleção.  |
| 3.5. Destino final*                           | Destino final da classe.   |
| 3.6. Natureza da intervenção                  | Indicação da natureza da intervenção, dono ou participante, da entidade produtora do processo de negócio.  |
| 3.7. Dono do PN                               | Indicação da entidade dona do processo de negócio. Informação a preencher apenas no caso de ter preenchido “participante” no campo acima.  |
| 3.8. Datas extremas                           | Datas extremas da informação / documentação a eliminar.  |
| 3.9. Número de agregações*                    | <p>Número das agregações identificadas nesta classe. Elemento de informação que resulta da soma das agregações identificadas na <i>Zona de identificação e controlo das agregações</i> dependente.</p>   |

Plataforma CLAV: contributo para a disponibilização de dados abertos  
da Administração Pública em Portugal

| Zonas e Elementos de informação                   | Nota explicativa  |
|---|---|
| 3.10. Dimensão por tipo de suporte                | Dimensão dos vários tipos de suporte da documentação / informação identificadas neste bloco.  |
| 4.Zona de identificação e controlo das agregações | Zona em que se individualiza as agregações dependentes de cada uma das classes identificadas na zona 3.   |
| 4.1. N.º da agregação                             | Identificador da agregação a eliminar.  |
| 4.2. Título da agregação                          | Título da agregação a eliminar.   |
| 4.3. Data de início de contagem do PCA            | Data de início da contagem do PCA. Corresponde à data de encerramento da agregação, de acordo com a forma de contagem estabelecida para a classe, e que marca o início de contagem do PCA. Esta data pode ser diferente da data de conclusão do procedimento, dado que tem de corresponder à especificidade da forma de contagem. |
| 4.4. Natureza da intervenção                      | Identificação da natureza da intervenção, dono do PN ou participante da entidade produtora no caso representado (aplica-se quando a entidade no âmbito de um PN pode intervir como dono numas circunstâncias e noutras como participante).  |

\* Informação preenchida automaticamente quando o auto de eliminação é efetuado a partir da Plataforma CLAV.

**Quadro 1**

Elementos informativos do auto de eliminação

Legislação: Catálogo de diplomas de suporte à LC e às tabelas de seleção. Este catálogo é gerido de forma incremental.

Entidades: Entidades da AP que intervêm no PN na qualidade de dono ou de participante, com consequências principalmente ao nível da conservação dos documentos.

Tipologia de Entidades: Para facilitar a gestão, dentro da Plataforma, das relações entre organizações e processos introduziu-se mais este conceito que é um agregador de organizações que partilham determinadas características. Exemplos: Agrupamentos de Centros de Saúde, Cartórios Notariais.

Foram ainda consideradas as seguintes entidades abstratas de gestão:

Utilizador: Este conceito representa os utilizadores da Plataforma que pertencem a um determinado perfil que determina o seu nível de permissões e consequentemente as operações e funcionalidades a que têm acesso.

Logs: Todas as ações realizadas na Plataforma ficam inscritas num registo de logs. Este funciona como um histórico de operações e serve de suporte a estatísticas de utilização e até

mesmo de responsabilização pelas operações realizadas. Nesse registo fica guardada a data, o tipo de operação e o utilizador que a realizou.

Foram definidos como requisitos de gestão as operações de manutenção e de relação entre as classes abstratas acima identificadas.

Como forma de complementar os requisitos, foram ainda identificadas questões a que a Plataforma CLAV deve responder, podendo estas ser transversais ou relativas às diversas relações estabelecidas entre as instâncias das diversas entidades da ontologia.

## 2) Requisitos para a interface e a interoperabilidade

Para garantir a interoperabilidade, o Projeto necessitou de efetuar o alinhamento com o Regulamento Nacional de Interoperabilidade Digital (RNID), aprovado pela RCM n.º 2/2018, que define as especificações técnicas e formatos digitais a adotar pela AP.

Considerou-se que, neste contexto, se aplicam os seguintes domínios do RNID:

- a) Formatos de dados, incluindo códigos de caracteres (neste caso, o formato Unicode UTF-8), formatos de som e imagens (fixas e animadas), audiovisuais, dados gráficos e de pré-impressão;
- b) Formatos de documentos (estruturados e não estruturados) e gestão de conteúdos, incluindo gestão documental (por exemplo, para a documentação serão usados os formatos XML e PDF);
- c) Tecnologias de interface Web, incluindo acessibilidade, ergonomia, compatibilidade e integração de serviços (no caso, no que diz respeito a tecnologias Web a utilizar: HTML, CSS, JavaScript, XSLT, XML);
- d) Especificações técnicas e protocolos de integração, troca de dados e orquestração de processos de negócio na integração inter-organismos (por exemplo, para a troca de informação entre processos serão usados *Web services* baseados em SOAP e SAML).
- e) Especificações técnicas e protocolos de comunicação em redes informáticas;
- f) Especificações técnicas de segurança para redes, serviços, aplicações e documentos.

O Projeto teve ainda a preocupação de garantir a disponibilização de serviços via iAP, enquanto plataforma de integração transversal à Administração Pública, quer para consumo de *Web services* disponibilizados por outras entidades organizacionais, quer para disponibilização de *Web services* para outras entidades (exemplo: plataforma de dados abertos Dados.gov).

Em concreto, a integração com a iAP far-se-á colocando a Plataforma CLAV como fornecedora de serviços nessa plataforma. Neste momento, pode ser feita através de *Web services* SOAP assíncronos com utilização de WS-addressing. Nos próximos meses estará também disponível a integração via REST.

A integração com o portal Dados.gov far-se-á disponibilizando a ontologia da Plataforma CLAV (engloba a Lista Consolidada, o catálogo legislativo e o catálogo de entidades) em formato RDF ou CSV enriquecido (neste momento a articulação com a

plataforma pode fazer-se por envio do ficheiro, mas a CLAV tem a API de dados preparada para uma integração automática).

Foram ainda definidos, em fase inicial, os requisitos necessários para potenciar a integração da Plataforma CLAV no LOD ("Linked Open Data"), de modo a permitir, entre outros, a exportação de vocabulários ontológicos em Resource Description Framework (RDF). Foram também considerados requisitos que implicam a modelação de dados em SKOS ("Simple Knowledge Organization System") e OWL ("Ontology Web Language").

### 3) Requisitos para a autenticação

Quanto aos mecanismos de autenticação previstos, procedeu-se à identificação de requisitos para a utilização do Cartão de Cidadão e da Chave Móvel Digital, enquanto mecanismos de autenticação e identificação eletrónica através do Fornecedor de Autenticação. A iAP já tem na sua plataforma de integração um fornecedor de autenticação que presta este serviço.

### 4) Modelo ontológico

#### 1 / Justificação da escolha do modelo ontológico

Atendendo à natureza hierárquica de algumas entidades do modelo, como a Lista Consolidada e as Tabelas de Seleção, bem como à quantidade e complexidade de relações existentes entre as entidades abstratas, sobre as quais é necessário extrair informação, o modelo encontrado para dar solução ao problema foi o ontológico.

O modelo ontológico baseia-se em relações binárias muito simples: X relaciona-se com Y pela relação r. Estas relações são normalmente representadas por triplos: (X, r, Y) em que:

- X: é o sujeito da relação e corresponde normalmente à instância de uma classe;
- r: é a relação, também designada por predicado;
- Y: é o objeto da relação e pode corresponder à instância de uma classe ou de um valor dum tipo primitivo (inteiro, decimal, tempo, string, etc).

Quando Y corresponde a uma instância de classe dizemos que estamos perante uma relação ou "Object Property". Quando Y é um valor dizemos que estamos perante um atributo ou "Data Property".

Estes triplos têm grande capacidade e alcance para ações de modelação. No entanto, há que acautelar algumas situações, como aquelas em que é necessário recorrer a atributos compostos. Neste caso, o atributo não é um valor simples, mas um conjunto de valores (tuplo). Para resolver estas situações, foi criada uma nova classe na ontologia: Atributo Composto. As suas instâncias corresponderão a atributos compostos em que os vários valores que os

compõem aparecerão como atributos destas instâncias (casos das “Notas de aplicação” e das “Notas de exclusão” referentes às classes da LC).

## 2/ Representação gráfica do modelo ontológico base

Neste contexto definiu-se um modelo ontológico com as seguintes entidades abstratas principais e respetivas relações:

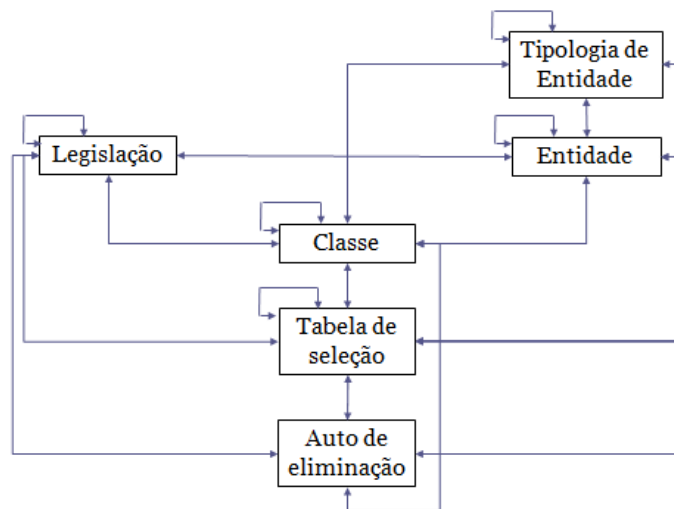


Figura 1

Modelo ontológico: principais entidades abstratas e respetivas relações

## 3/Modelo das Classes / Lista Consolidada

A realização de operações sobre a Lista Consolidada é controlada pelo nível de acesso de cada utilizador. Todas as operações que alteram o estado da LC só estão disponíveis para utilizadores com acesso de administração.

A LC, como já referido, é uma hierarquia de classes. Todas elas são constituídas pelos seguintes atributos:

**Código:** Código identificador atribuído pela DGLAB aquando da criação da classe;

**Título:** Atributo textual que contém o título pelo qual a classe será reconhecida;

**Descrição:** Atributo textual que contém uma indicação do âmbito e do que constitui a classe, como por exemplo o seu início, etapas e fim, no caso de processos de negócio;

**Notas de Aplicação:** Atributo textual que descreve em que situações deve ser usada a classe. Podem ser “genéricas” (comuns a todas as organizações) ou específicas (próprias de uma ou mais organizações). No caso de serem específicas, terão de ter uma relação com as organizações em causa. Por este motivo, este atributo foi promovido a atributo composto;

**Notas de Exclusão:** Atributo textual que descreve em que situações a classe não deverá ser usada. Podem ser “genéricas” (comuns a todas as organizações) ou específicas (próprias de uma ou mais organizações). Podem ainda remeter para outras classes. Nesse caso, o texto descritivo irá conter uma ou várias referências que se pretendem

transformar em hiperligações ativas na interface da Plataforma. Devido a poderem ser específicas e, também, poderem remeter para outras classes, este atributo foi também promovido a atributo composto.

As classes de nível 3, relativas a processos de negócio, têm mais dois atributos:

Tipo de Processo: Pode ser comum ou específico;

Processo Transversal: Sim ou Não (se “Sim”, intervirão como participantes no processo outras organizações).

As classes de nível 3 ou de nível 4 (quando estas existirem) têm ainda os seguintes atributos adicionais:

Prazo de Conservação Administrativa (PCA): prazo durante o qual o documento deve ser preservado (pode ser alterado ao longo do tempo, mas todos os valores terão de ser guardados e associados às datas em que estavam em vigor). Devido à sua estrutura, este atributo foi também promovido a atributo composto;

Justificação do Prazo de Conservação: fundamentação do prazo de conservação indicado. Pode decorrer do tipo de relação com a entidade Legislação ou do tipo de relação entre processos de negócio (relação suplementar). Tem relação com a Legislação em 90% dos casos e tem de estar associada ao fator tempo. Devido à sua estrutura, este atributo foi também promovido a atributo composto;

Destino Final (DF): apresenta um valor fixo – conservação, eliminação e conservação parcial. Neste último caso, é preciso guardar a percentagem do que se conserva. Devido à sua estrutura, este atributo foi também promovido a atributo composto;

Justificação do Destino Final: fundamentação para o destino final; pode decorrer do tipo de relação com a entidade Legislação ou do tipo de relação entre processos de negócio (relação de síntese ou relação complementar);

No modelo, quando um processo de terceiro nível se subdivide em quartos níveis, os atributos PCA, Justificação do PCA, DF e Justificação do DF apenas estão presentes ao nível mais baixo.

Na criação de uma nova classe devem ser observados os seguintes requisitos:

- O código deverá ser único e seguir o formato pré-definido e que pode expressar numa das expressões regulares seguintes:

`/^[0-9]{3}$/`

`/^[0-9]{3}\.[0-9]{2}$/`

`/^[0-9]{3}\.[0-9]{2}\.[0-9]{3}$/`

`/^[0-9]{3}\.[0-9]{2}\.[0-9]{3}\.[0-9]{3}$/`

- A classe terá um tempo de vida ativa, o instante 0 coincidirá com a data em que se torna ativa e o seu período de vida termina quando a operação de a inativar for executada; ambas as datas devem ser registadas na sua informação pois condicionarão os resultados das pesquisas e das queries;
- As classes de 1.º e 2.º nível têm apenas informação descritiva;

- A criação de classes de 3.º nível segue um algoritmo mais complexo: a classe de 3.º nível pode desdobrar-se em classes de 4.º nível.

Foram também identificados os seguintes tipos de relações entre classes de 3.º nível:

#### 1/ Relações simétricas

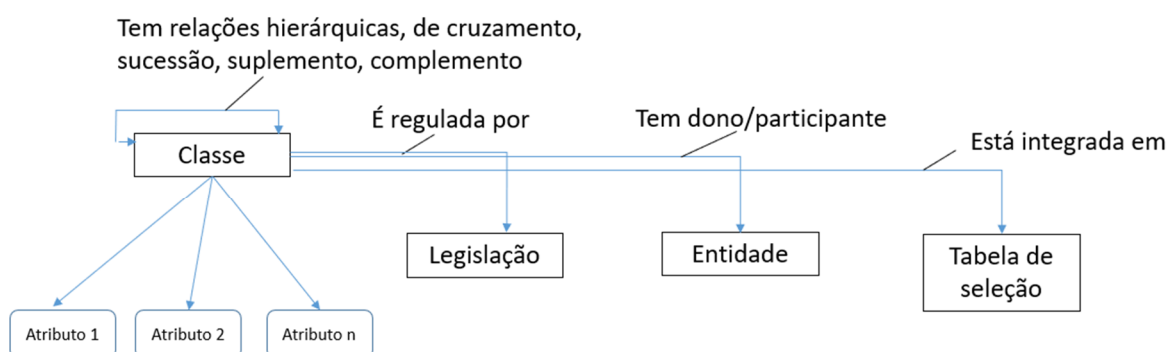
- Complementar – quando dois processos, decorrendo de forma paralela, adicionam um ao outro informação complementar;
- Cruzada – quando existe interseção de dois processos em determinado momento, seguindo cada um percursos distintos.

#### 2/ Relações assimétricas

- Síntese – quando um processo condensa a informação de outro processo (ou quando uma etapa de um processo condensa outras etapas desse processo). A relação está indicada como “Sintetiza ou éSintetizado”;
- Sucessão – quando o produto de um processo dá origem a outro processo (o precedente não existe sem o anterior). A relação está indicada como “éSucessorDe” ou “éAntecessorDe”;
- Suplementar – quando um PN recolhe e analisa a informação contida noutros PN cotejando-os entre si, não lhes adicionando conteúdo informativo. A relação está indicada como “éSuplementoDe” ou “éSuplementoPara” no processo relacionado.

#### 3/ Donos e Participantes no processo

Neste contexto, foi definido um modelo ontológico para as classes que integram a Lista Consolidada, com a seguinte configuração:



**Figura 2**

Modelo ontológico aplicado às classes de suporte à LC

#### 4/ Modelo das Tabelas de Seleção



Uma Tabela de Seleção (TS) é um subconjunto da LC, e por isso segue o mesmo modelo ao qual acrescem alguns atributos adicionais:

ResponsávelDF: identificação do responsável pelo cumprimento do destino final;

Status: a TS pode estar num de 4 estados possíveis:

1. Proposta,
2. Validada pela DGLAB,
3. Publicada,
4. Revogada.

Acresce ainda os atributos e relações relativos à identificação do diploma de suporte à TS.

Foi também definido um modelo ontológico para as classes que integram as tabelas de seleção, com a seguinte configuração:

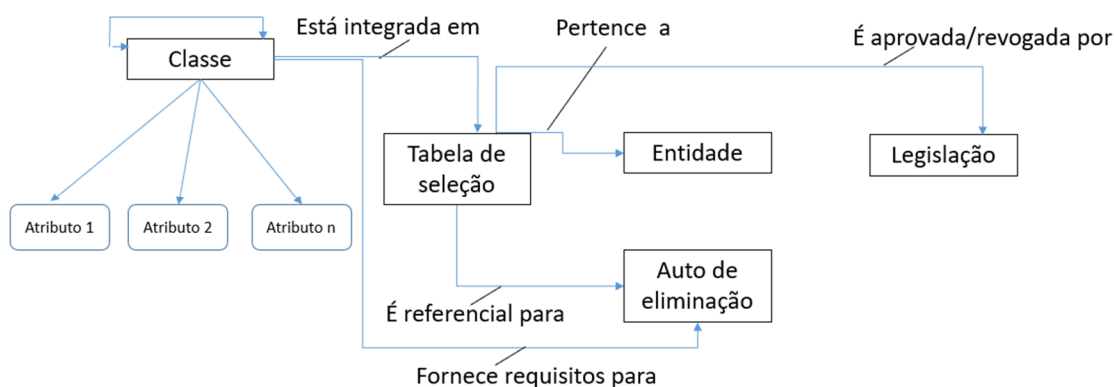


Figura 3

Modelo ontológico aplicado à Tabela de seleção

## 5/ Modelo da Legislação

O modelo é constituído pelos seguintes atributos:

Código: código identificativo do diploma legal. Pode ser uma forma abreviada desde que se garanta que não há sobreposições entre os códigos;

Sumário: atributo textual que contém o título do elemento legislativo e que, nalguns casos, descreve o âmbito e conteúdo;

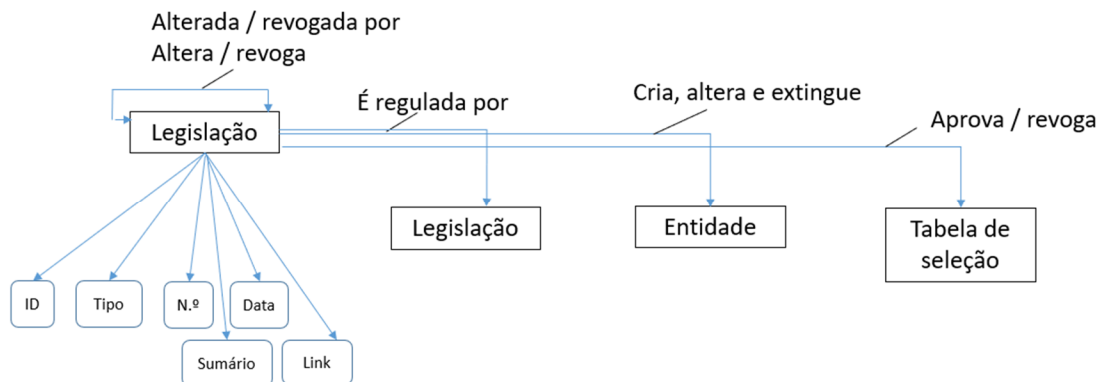
Data: atributo que contém a data do elemento legislativo (entrada em vigor);

Tipo: atributo textual que contém o tipo do elemento legislativo;

Número: atributo textual que contém o número do elemento legislativo;

Link: atributo textual que contém o endereço Web para o documento original.

Foi ainda definido um modelo ontológico para as classes de suporte à Legislação com a seguinte configuração:



**Figura 4**  
Modelo ontológico aplicado à Legislação

## 6/ Modelo da Entidade

O modelo é constituído pelos seguintes atributos:

Código: código identificativo da entidade organizacional. Para garantir que não há sobreposições entre os códigos, optou-se por um código sequencial numérico, em alternativa à sigla da entidade, atendendo a que esta pode ocorrer de forma repetida em diferentes entidades, em distintos períodos cronológicos;

Nome: atributo textual que contém o nome designativo da entidade;

Sigla: atributo textual que contém a sigla da entidade.

Foi também definido um modelo ontológico para as classes de suporte às entidades com a seguinte configuração:

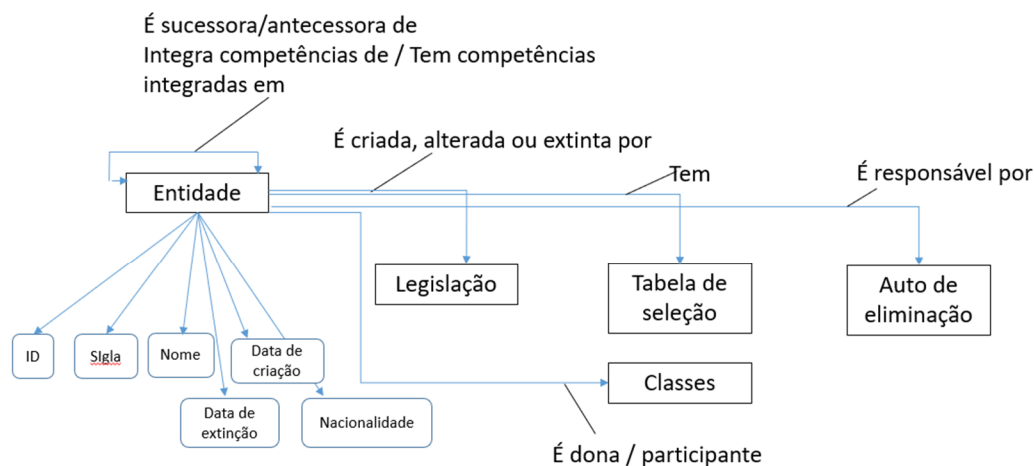


Figura 5

Modelo ontológico aplicado às Entidades organizacionais

## 7/ Modelo da Tipologia de Entidade

O modelo é constituído pelos seguintes atributos:

**Código:** código identificativo da tipologia de entidade organizacional. Pode ser uma sigla desde que se garanta que não há sobreposições entre os códigos;

**Nome:** atributo textual que contém o nome designativo da tipologia de entidade;

**Descrição:** atributo textual opcional que contém uma memória descritiva e resumida da tipologia de entidade;

**Entidades:** lista de entidade enquadradas nesta tipologia (opcional).

## 5) Especificação da ontologia

A *Ontology Web Language* (OWL) é uma linguagem para a especificação formal de conhecimento (ontologias) que está definida em várias versões permitindo criar especificações com vários níveis de complexidade. As ontologias são uma maneira formal de descrever taxonomias e sistemas de classificação, definindo essencialmente a estrutura do conhecimento para vários domínios: os substantivos, que representam classes de objetos e os verbos, que representam as relações entre os objetos. As ontologias são tipicamente muito flexíveis, uma vez que se destinam a representar informações provenientes de todos os tipos de fontes de dados heterogéneas. Uma solução tecnológica baseada em OWL, permite um crescimento incremental do modelo e dá uma liberdade muito grande para se poderem fazer alterações mesmo no decurso do desenvolvimento e este é um ponto fundamental no desenvolvimento da CLAV.

Para o desenvolvimento da ontologia, no Projeto CLAV seguiu-se uma metodologia iterativa: começou-se com uma versão básica à qual se foram juntando elementos, à medida que se foram detalhando os requisitos.

## 1 / Representação gráfica da informação presente na ontologia

Para a exploração visual e para validação do trabalho de modelação recorreu-se principalmente ao Visual graph, como se pode ver no exemplo da Figura 6.

Neste exemplo a rede encontra-se centrada na classe 150.10.700. À esquerda pode ver-se o grafo de relações desta classe e, à direita, o painel informativo com os seus atributos.

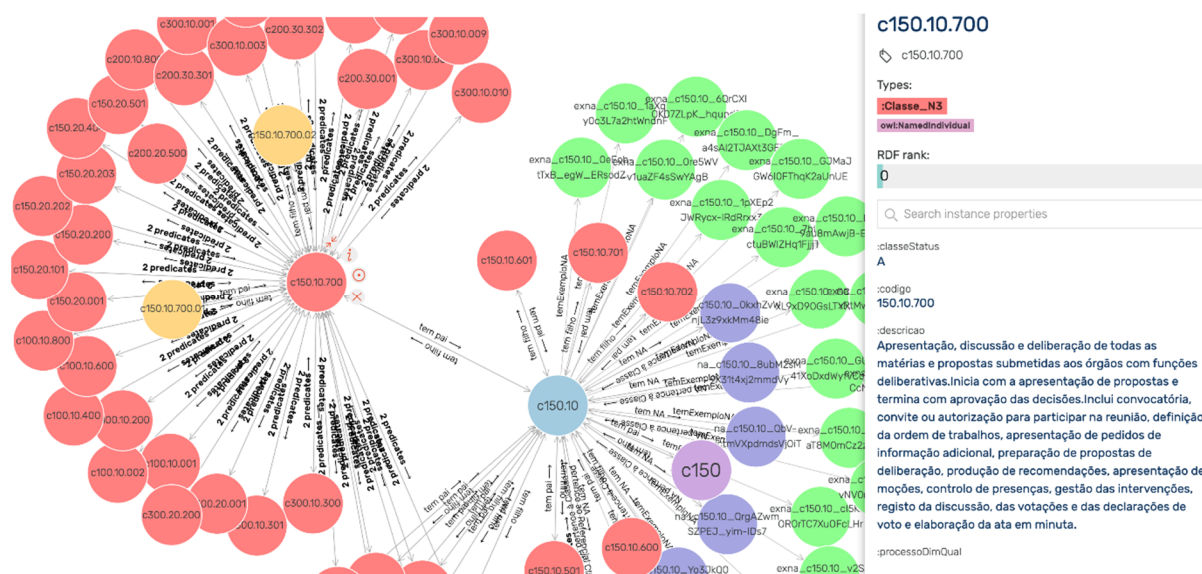


Figura 6

Classe 150.10.700, seus atributos e relações

## 6) Resultados da disponibilização da ontologia

A migração de dados para a Plataforma CLAV, para que a implementação da ontologia se pudesse vir a realizar, foi antecida de um trabalho exaustivo de revisão e de normalização da informação da Lista Consolidada (LC), que se encontrava em folha de cálculo. Para além da normalização do preenchimento de todos os campos, que foi posta em prática desde a criação da LC, foram ainda criadas listas de vocabulário controlado aplicadas a alguns dos campos. Destaca-se o resultado do trabalho de normalização da “Forma de Contagem do Prazo de Conservação Administrativa”, que contava com mais de 300 expressões diferentes, o qual deu origem a uma lista de vocabulário controlado com apenas 7 entradas.

A primeira versão da Plataforma CLAV foi disponibilizada ao público a 22 de julho de 2018, no endereço: <http://clav.dglab.gov.pt>.

No processo de consulta, o utilizador pode efetuar a ligação entre as várias entidades da ontologia, consoante as suas necessidades (por exemplo, pode verificar que entidades organizacionais são participantes num determinado processo de negócio selecionado ou que legislação o regula e a seguir pode conhecer que outras classes deverão constar na sua tabela

de seleção).



The screenshot displays the CLAV platform interface. At the top, there is a navigation menu with options like 'Lista Consolidada', 'Tabelas de Seleção', 'Autos de Eliminação', 'Entidades', 'Tipologias de Entidades', 'Legislação', 'Estatísticas', and 'Termos de Índice'. A search bar is located on the left side. The main content area shows a search for the class '150.10.700 - Reunião de órgãos deliberativos'. The interface includes a tree view on the left showing a hierarchy of classes. The detailed view on the right shows the following information:

- Estado:** Ativo
- Classe Pai:** 150.10 - Definição e avaliação de políticas
- Descendência:**
  - 150.10.700.01 - Reunião de órgãos deliberativos: preparação
  - 150.10.700.02 - Reunião de órgãos deliberativos: realização
- Código:** 150.10.700
- Título:** Reunião de órgãos deliberativos

The 'Descritivo da Classe' section provides a detailed description: 'Apresentação, discussão e deliberação de todas as matérias e propostas submetidas aos órgãos com funções deliberativas. Inicia com a apresentação de propostas e termina com aprovação das decisões. Inclui convocatória, convite ou autorização para participar na reunião, definição da ordem de trabalhos, apresentação de pedidos de informação adicional, preparação de propostas de deliberação, produção de recomendações, apresentação de moções, controlo de presenças, gestão das intervenções, registo da discussão, das votações e das declarações de voto e elaboração da ata em minuta.'

Under 'Nota(s) de Aplicação:', the following are listed:

- Reuniões ordinárias, extraordinárias e descentralizadas
- Reuniões do Conselho de Ministros
- Reuniões dos Secretários de Estado
- Reuniões de Assembleia Intermunicipal
- Reuniões de Assembleia de Freguesia

Figura 7

Plataforma CLAV: consulta da classe 150.10.700 (versão disponibilizada em 2018)

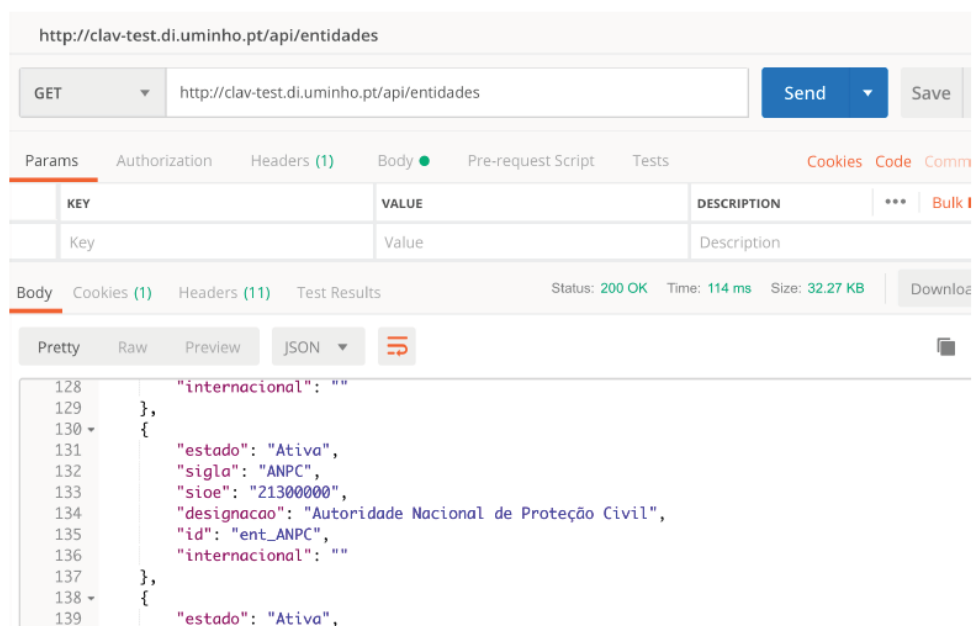
Os trabalhos desenvolvidos potenciam, como referido, a consulta e disponibilização de dados abertos. Em concreto, os catálogos de dados abertos (com licença CC0 (Creative Commons CC0 1.0 Universal Public Domain Dedication) disponíveis, abrangem a Lista Consolidada, e proximamente as tabelas de seleção das várias entidades do Estado, dados sistematizados sobre a eliminação de informação pública e dados estatísticos sobre a utilização da Plataforma que contribuirão para a monitorização da aplicação do RJCAIA.

O modelo de estruturação dos dados, que aprofundámos anteriormente, permite que estes possam ser descarregados a partir da Plataforma em CSV, RDF, OWL, SKOS e XML.

Como já referimos, esta estruturação potencia o uso dos dados através de tecnologias associadas à Web semântica, entre outras. O facto de os dados terem sido guardados numa base de dados de triplos RDF e o respetivo modelo ter sido especificado em OWL permite a sua exploração com linguagens como o SPARQL ("SPARQL Protocol and RDF Query Language") e a sua integração com o LOD ou com o Portal Dados.gov. Para estes fins, estão planeadas exportações em RDF, SKOS e OWL.

Foram ainda criados métodos que permitem que processos ou aplicações possam aceder aos dados sem intervenção humana. Foi desenvolvida uma API de dados (conjunto de funções e procedimentos que permitem a criação de aplicações que podem aceder à informação da base de dados ontológica da CLAV) para funcionar como interface máquina-máquina. Esta interface foi desenvolvida seguindo uma metodologia REST. As operações estão acessíveis através de URLs e respondem com informação em JSON (em breve serão

disponibilizados outros formatos de exportação ao nível da API).



**Figura 8**

Plataforma CLAV: teste de utilização de API (consulta de Entidades)

Uma ideia subjacente a esta arquitetura aplicacional é, no futuro, a possibilidade de criação de uma série de novas aplicações usando a API da CLAV com lógicas de negócio distintas da planeada originalmente. Por exemplo, navegadores gráficos, programas que analisem a similaridade semântica de processos de negócio e organizações, integração com sistemas de gestão documental, aplicações de análise estatística e outros.

Para potenciar a sua estratégia de disponibilização de dados abertos, a DGLAB tenciona ainda alargar a divulgação da Plataforma e do modo como se pode potenciar o uso deste tipo de dados, estando programadas sessões nacionais, regionais e especializadas, orientadas a públicos promotores do seu uso, onde se destaca a Administração Pública, o meio académico e o meio empresarial, para além de associações cívicas preocupadas com a transparência da atuação da AP.

## 7) Trabalhos desenvolvidos em 2019

As funcionalidades de atualização das classes, entidades organizacionais e legislação relativas à Lista Consolidada bem como elaboração de tabelas de seleção e submissão de autos de eliminação, foram disponibilizadas no início do segundo semestre de 2019, numa nova versão da Plataforma.

Constataram-se ainda as seguintes alterações, em 2019:

1. A arquitetura aplicacional evoluiu para uma arquitetura baseada em microserviços, tendo-se isolado cada componente num “container” Docker que pode ser descarregado, instalado e correr individualmente. Neste momento, a Plataforma está dividida em quatro microserviços: um contendo o servidor GraphDB, que guarda e suporta a base de dados ontológica, um contendo o servidor MongoDB, que guarda e suporta a base de dados de gestão da Plataforma, um contendo a API de dados e outro contendo a interface Web. Atendendo a que a API de dados necessita dos servidores de bases de dados para funcionar, pretende-se dividir a aplicação em dois blocos funcionais, a API de dados e o “front-end” Web. Cada um destes pode ser instalado em máquinas diferentes, podendo até ser replicado, se for necessário, por motivos de eficiência.
2. A API de dados foi melhorada e irá apresentar duas interfaces: uma de consumo interno para o “front-end”, com todas as funcionalidades disponíveis, e uma interface aberta ao utilizador externo, apenas com as funcionalidades de consulta, pesquisa e recuperação de informação. A API de dados, nas duas versões, tem o acesso protegido por token temporal que é obtido através de uma ação de registo simples. A atividade associada a cada um dos tokens emitidos é monitorizada podendo ajudar no balanceamento de carga nos servidores, na otimização de operações e na deteção do uso indevido da API. O “front-end” Web já permite, de forma assistida, a criação de tabelas de seleção, garantindo que as mesmas são coerentes e completas. Para além disso, permite consultar toda a informação presente na Plataforma bem como submeter pedidos de alteração ou criação dos vários itens de informação.
3. Foram definidos XML Schema para os processos de negócio. A API de dados, que em breve será disponibilizada publicamente, está a usá-los na exportação de dados em XML.
4. Foi ainda definido um formato CSV para importação e exportação de informação da Plataforma, que também será disponibilizado.

Como referimos, o modelo principal foi especificado em OWL, tendo sido materializado numa ontologia. Devido à complexidade deste modelo, principalmente, a grande quantidade de conhecimento implícito que foi necessário tornar explícito na ontologia, fez-se um trabalho de identificação de invariantes sobre o modelo. Foram inventariados cerca de 38 invariantes, alguns dos quais com ramificações. Foram todos especificados formalmente e traduzidos em queries que podem agora ser usadas para a construção de um validador autónomo ou integrado na Plataforma CLAV, que detetará

anomalias na informação que venha a ser introduzida na Plataforma (muitos dos invariantes exibem um grau de complexidade elevado e com dependências em muitas partes do modelo, o que torna a sua verificação por humanos quase impossível).

## Discussão

A construção do referencial Lista Consolidada que está na base do Projeto CLAV é um exemplo de resiliência face aos múltiplos constrangimentos derivados das alterações político-administrativas, e da natural resistência à mudança por parte das organizações públicas e dos seus profissionais. Os princípios e valores associados ao referencial e ao Projeto têm-se afirmado como uma vantagem em contexto de transformação, reforçando os benefícios do trabalho colaborativo e da abordagem funcional e suprainstitucional.

A Plataforma CLAV e a proposta de novo Regime jurídico para a classificação e avaliação da informação arquivística, que se prevê que possa alavancar a sua utilização, são suportes indispensáveis para a publicitação da atuação do Estado, contribuindo para a transparência administrativa e para uma cidadania ativa. Consultando os dados da ontologia, qualquer cidadão pode conhecer o que o Estado faz e controlar, por exemplo, que tipo de informação é capturada e salvaguardada, por quanto tempo e qual é legalmente eliminada.

A disponibilização dos dados abertos da ontologia permitirá ainda outros usos com algum impacto, quer na atividade da própria Administração Pública, quer das empresas, a exemplo daquelas que prestam consultoria ao Estado no domínio da gestão da informação. No primeiro caso, os dados abertos disponibilizados a partir da Plataforma permitirão a inclusão atualizada e direta em sistemas de gestão documental facilitando uma maior interoperabilidade semântica na AP. Ao mesmo tempo possibilitam uma melhor organização, avaliação e recuperação da informação, evitando a acumulação de informação descontrolada na Administração e contribuindo ainda, através do princípio da corresponsabilização, para que a entidade dona dos processos de negócio possa conservar a informação com destino final de conservação permanente e salvaguardá-la a longo prazo através de medidas de preservação digital adequadas. No caso das empresas de consultoria, esta disponibilização de dados abertos permitirá, por exemplo, que estas se possam focar na aplicação de instrumentos de gestão de documentos (por exemplo, tabelas de seleção) mais do que na sua produção.

A divulgação da Lista Consolidada e das tabelas de seleção através de dados abertos permitirá ainda uma maior conformidade da aplicação dos requisitos de gestão da informação com o Regulamento Geral da Proteção de Dados (RGPD). Estes instrumentos apoiam a contextualização dos dados pessoais, no âmbito dos respetivos processos de negócio, indicam a finalidade para o seu tratamento e a fundamentação para a sua eliminação ou conservação. Note-se que a nova lei de execução em Portugal do RGPD – Lei n.º 58/2019, de 8 de agosto, clarifica a articulação dos prazos de conservação com o atual regime jurídico da avaliação, baseado no Decreto-Lei n.º 447/88. No seu artigo 21.º refere-se expressamente



que “o prazo de conservação de dados pessoais é o que estiver fixado por norma legal ou regulamentar”, remetendo para o contexto de aprovação destes prazos e destinos finais através de portaria de gestão de documentos, tal como previsto no referido Decreto-Lei.

## Conclusão

Nos últimos anos, a Direção-Geral do Livro, dos Arquivos e das Bibliotecas (DGLAB) desenvolveu vários projetos com vista à construção de um referencial para a classificação e avaliação da informação do Estado, que facilitasse a elaboração de tabelas de seleção pelos organismos públicos e tornasse mais eficiente e económica a gestão documental. A atual Lista Consolidada para a classificação e avaliação da informação pública (LC) é tributária dos referidos projetos e dos seus resultados.

Desde o primeiro momento que se tornou óbvio que era necessário disponibilizar os dados da Lista Consolidada em acesso aberto, o que a DGLAB concretiza atualmente através da Plataforma CLAV, que possui várias funcionalidades de consulta já disponíveis desde julho de 2018.

Os trabalhos desenvolvidos e os que se encontram em curso demonstram uma grande preocupação pela construção de um modelo ontológico, explicitado nesta comunicação, que está na base da estrutura dos dados disponibilizados em acesso aberto através da Plataforma, com vista ao fomento da transparência, de uma Administração aberta e eficiente e de uma maior reutilização da informação pública. Estão previstas ainda outras formas de disponibilização deste tipo de dados, a exemplo da plataforma governamental Dados.gov.

Deste modo, a Plataforma CLAV assume-se como um instrumento fundamental para a concretização das políticas públicas de modernização administrativa, de inovação, de informação e de acesso aberto.

## Referências bibliográficas

AGÊNCIA PARA A MODERNIZAÇÃO ADMINISTRATIVA (2016) – *Guia de introdução aos dados abertos*. [Em linha]. Lisboa: Agência para a Modernização Administrativa. [Consult. 12 Set. 2019]. Disponível na Internet: <URL: [https://www.ama.gov.pt/documents/24077/24804/guia\\_introdu\\_o\\_dados\\_abertos\\_ama.pdf/9b40b98c-4935-471b-af5d-f6f6a656edc0](https://www.ama.gov.pt/documents/24077/24804/guia_introdu_o_dados_abertos_ama.pdf/9b40b98c-4935-471b-af5d-f6f6a656edc0) >.

LOURENÇO, A.; RAMALHO, J. C.; GAGO, M. R.; PENTEADO, P. (2018) – Transformação digital: novas políticas e procedimentos para a classificação e a avaliação da informação. 13.º Congresso Nacional de Bibliotecários, Arquivistas e Documentalistas. Atas. [Em linha]. [Consult. 12 Set. 2019]. Disponível na Internet: <URL: <https://www.bad.pt/publicacoes/index.php/congressosbad/article/view/1861> >.

LOURENÇO, A.; RAMALHO, J. C.; GAGO, M. R.; PENTEADO, P. (2017) – Plataforma M51-CLAV: o que há de novo? [Em linha]. *CIGIA – Conferências Internacionais de Gestão da Informação e Arquivos*. [Em linha]. [Consult. 12 Set. 2019]. Disponível na Internet: <URL: [https://www.bad.pt/eventos/wp-content/uploads/2018/01/CIGIA\\_COM\\_01.pdf](https://www.bad.pt/eventos/wp-content/uploads/2018/01/CIGIA_COM_01.pdf) >.

LOURENÇO, A.; RAMALHO, J. C.; GAGO, M. R.; PENTEADO, P. (2016) – Simplex +: o que precisamos para além da Medida 51? [Em linha]. *Actas do Encontro Nacional de Arquivos Municipais*. N.º 12. [Consult. 12 Set. 2019]. Disponível na Internet: <URL: <https://www.bad.pt/publicacoes/index.php/arquivosmunicipais/article/view/1535/1465> >.

LOURENÇO, A.; PENTEADO, P. (2016) – *Avaliação Suprainstitucional da Informação Arquivística (ASIA): documento metodológico*. [Em linha]. Lisboa: DGLAB. [Consult. 12 Set. 2019]. Disponível na Internet: <URL: [http://arquivos.dglab.gov.pt/wp-content/uploads/sites/16/2016/03/ASIA\\_Doc-metodologico2016-03-10.pdf](http://arquivos.dglab.gov.pt/wp-content/uploads/sites/16/2016/03/ASIA_Doc-metodologico2016-03-10.pdf) >.

LOURENÇO, A.; PENTEADO, P. (2015) – A caminho da ASIA – Avaliação Suprainstitucional da Informação Arquivística. *Atas do 12.º Congresso Nacional de Bibliotecários, Arquivistas e Documentalistas*. [Em linha]. [Consult. 12 Set. 2019]. Disponível na Internet: <URL: [http://www.bad.pt/publicacoes/index.php/congressosbad/article/view/1458/pdf\\_90](http://www.bad.pt/publicacoes/index.php/congressosbad/article/view/1458/pdf_90) >.

LOURENÇO, A.; PENTEADO, P. (2014) – Una estrategia para mejorar el acceso y la reutilización de la información pública en Portugal: el papel de la interoperabilidad semántica. *Girona 2014: Archivos e Industrias Culturales*. [Em linha]. [Consult. 12 Set. 2019]. Disponível na Internet: <URL: <http://www.girona.cat/web/ica2014/ponents/textos/id200.pdf> >.

MUNIR, K.; ANJUM, M. S. (2018) – The use of ontologies for effective knowledge modelling and information retrieval. *Applied Computing and Informatics*. 14, pp. 116–126. [Em linha]. [Consult. 12 Set. 2019]. Disponível na Internet: <URL: <https://reader.elsevier.com/reader/sd/8DF0CA070CBECDABFCF8445C3216C3CAD8AB256F2A73440491CA129973A1B90C4C3C4A19AF3B5390C535E3020E9D48EF> >.

SASHINSKAYA, M. (2017) – *Open data: all you want to know about open data (Big data, transparency, urbanism, transportation, sustainable cities, innovations, smart governance, e-government)*. S.l.: CreateSpace Independent Publishing Platform.

SILVA, Paulo Pereira da (2017) – *As políticas de Open Data em Portugal: análise da sua implementação e impacto*. [Em linha]. Coimbra: Universidade de Coimbra. Tese de mestrado. [Consult. 12 Set. 2019]. Disponível na Internet: <URL: [estudogeral.sib.uc.pt/handle/10316/47025](http://estudogeral.sib.uc.pt/handle/10316/47025) >.

SHINGH, J. (2017) – *Open Data 101: The latest trends, challenges and research in government open data*. [Erindale, Australian Capital Territory] [Cooee Press].