

Del sistema de información 2ArchIS al DataRepositóriUM: el estudio de caso de hallazgos arqueológicos

From the archaeological information system 2ArchIS to the DataRepositóriUM: case study for archaeological findings

RECIBIDO: 25 DE MAYO DE 2020
ACEPTADO: 22 DE JUNIO DE 2020

NATÁLIA BOTICA

Unidade de Arqueologia da Universidade do Minho, Lab2PT.

E-mail: nb@uaum.uminho.pt

Orcid ID: <http://orcid.org/0000-0003-1080-4785>

FERNANDA MAGALHÃES

Unidade de Arqueologia da Universidade do Minho, Lab2PT.

E-mail: fmagalhaes@uaum.uminho.pt

Orcid ID: <http://orcid.org/0000-0002-2858-6432>

DIEGO MACHADO

Unidade de Arqueologia da Universidade do Minho.

E-mail: diegosfmachado@gmail.com

Orcid ID: <http://orcid.org/0000-0002-6144-4495>

LUÍS FONTES

Unidade de Arqueologia da Universidade do Minho, Lab2PT.

E-mail: lfontes@uaum.uminho.pt

Orcid ID: <http://orcid.org/0000-0002-3725-0797>

Resumen: Para promover la ciencia abierta y la reutilización de datos, es necesario que estos estén disponibles en repositorios abiertos que garanticen su accesibilidad y permanencia. Los repositorios de datos arqueológicos son particularmente importantes ya que muchos de estos conjuntos de datos son el resultado de actividades irrepetibles y ofrecen perspectivas casi infinitas para su reutilización. Sin embargo, no es suficiente transferir datos arqueológicos a repositorios para su difusión y preservación. Para que los datos sean Localizables, Accesibles y Reutilizables, deben obedecer las directrices que aseguran el uso de un esquema de metadatos adecuado, identificadores persistentes, vocabularios bien definidos, procedimientos para normalizar y mejorar la calidad de los datos y formatos de archivo sostenibles. Para lograr este objetivo, presentamos

la metodología utilizada para registrar datos arqueológicos en el contexto del trabajo de campo y de oficina, resultantes del trabajo realizado por la Unidade de Arqueología de la Universidade de Minho. Estos conjuntos de datos se registran en el sistema de información 2ArchIS y se exportan para su disponibilidad en el repositorio científico «DataRepositóriUM», de acceso abierto y asociados con metadatos que promueven su permanencia, difusión y reutilización por parte de múltiples usuarios, en varios proyectos de investigación. También presentamos algunos ejemplos de trabajos de visualización e investigación realizados sobre estos conjuntos de datos que, al compartirlos en el repositorio, se pueden utilizar para reproducir la investigación ya realizada, pero también en nuevos trabajos de investigación.

CAUN 29 (2021): [1-16]

ISSN: 1133-1542. ISSN-e: 2387-1814

DOI: <http://doi.org/10.15581/012.29.005>

Palabras Clave: Arqueología, Base de datos, integración de datos, Repositorio de datos, *Bracara Augusta*.

Abstract: To promote open science and data reuse, it is necessary to put data in open repositories to ensure its accessibility and permanence. Archaeology data repositories are particularly important, as many of these data sets result from unrepeatable activities and offer almost infinite possibilities for reuse. However, it is not enough to transfer archaeological data to repositories for dissemination and preservation. For data to be Findable, Accessible and Reusable, it is necessary to follow guidelines that ensure the use of an appropriate metadata scheme, persistent identifiers, well-defined vocabu-

laries, procedures to normalize and improve data quality and sustainable file formats. In this work we present the methodology used to record the findings of the coin type and their contexts, as a result from the archaeological works carried out by the Archaeology Unit of the University of Minho and its availability, in open access, in the scientific repository DataRepositóriUM.

We are also presenting some examples of visualization and research works carried out on these data sets that, due to their sharing in the repository, can be used not only to reproduce the research that has already been carried out, but also as a support for new research works.

Keywords: Archaeology, Data Base, Data Integration, Data Repository, *Bracara Augusta*.

1. INTRODUÇÃO

A O longo dos últimos anos, a Arqueologia tem produzido uma enorme quantidade de dados, sendo que uma parte considerável já nasce em formato digital. O registo dos dados arqueológicos deve cumprir duas funções, por um lado, viabilizar a preservação para memória futura, permitindo a sua utilização e análise e, por outro, possibilitar a disponibilização para serem reutilizados e sujeitos a novas análises, sob outras perspetivas ou com novas ferramentas. Neste sentido, a investigação em Arqueologia, que tem como ponto de partida a pesquisa documental, bibliográfica e de registos de dados, tem-se revelado uma tarefa difícil, demorada e incompleta. Se é certo que apenas alguma desta informação está publicada *online*, a que está difundida ou não é facilmente localizável ou não está acessível, existindo uma perceção geral de que o nível de partilha e reutilização de dados em arqueologia é muito baixo (Lodwick, 2019).

Embora a Arqueologia incorpore nos seus procedimentos de trabalho e de registo muitos dados que são nativamente digitais, os dados das escavações arqueológicas, que resultam de ações irrepetíveis, ou não estão acessíveis ou estão plasmados em relatórios e publicações, de difícil acesso e reutilização. As tecnologias de informação criaram uma oportunidade para tornar acessíveis os arquivos de escavação e muitos países têm investido em tecnologias, infraestruturas e normas para digitalização, preservação e disseminação do conhecimento arqueológico (Laužikas *et al.*, 2018). A Direcção-Geral do Património Cultural (DGPC), organismo que em Portugal tutela toda a atividade arqueológica, criou um Sistema de Informação e Gestão Arqueológica «Endovélico» e que constitui o principal instrumento de gestão da atividade arqueológica e de armazenamento de informação a nível nacional. Os conteúdos do

«Endovélico» são disponibilizados na internet no «Portal do Arqueólogo». No entanto, se por um lado este sistema deixa de fora muita informação produzida nas rospecções, escavações arqueológicas, trabalhos de gabinete e em projetos de investigação, por outro lado, são conteúdos maioritariamente partilhados em arquivos PDF, não estruturados, onde os conjuntos de dados são dificilmente localizáveis e acessíveis. É, portanto, urgente que seja criado um repositório de dados arqueológicos que aloje e torne acessíveis e persistentes os conjuntos de dados que são diariamente produzidos, quer pelos trabalhos arqueológicos, quer pelos projetos de investigação.

Os projetos de investigação financiados, quer a nível nacional quer pela Comissão Europeia, impõem agora aos investigadores a obrigação de apresentação de Planos de Gestão de Dados e de proceder ao depósito de dados e de resultados da investigação em repositórios persistentes. No entanto, apenas um número reduzido de países tem repositórios de dados arqueológicos, onde esses conjuntos podem ser depositados e acessíveis de forma livre e aberta e disponíveis para reutilização. O projeto ARIADNE e ARIADNEPlus, um dos maiores projetos internacionais, tem como principal missão a criação de uma infraestrutura europeia de pesquisa on-line para dados arqueológicos. Este projeto, que envolve 23 países parceiros, combina Arqueologia e Ciência da Computação para enfrentar os desafios de lidar com diferentes escolas e práticas arqueológicas, diferentes condicionalismos legais e diferentes capacidades de preparação e gestão de dados. Este projeto evidenciou a falta de equidade dos países participantes, onde apenas um número reduzido de países tem repositórios nacionais para dados arqueológicos, como o ADS (Reino Unido), DANS (Países Baixos) e SND (Suécia), e, na ausência destes repositórios, os dados estão dispersos, sob pena de se perderem (Štular, Corns, 2014). Este registo arqueológico disperso, que integra desde dados de escavação inseridos diretamente em equipamentos móveis, registo fotográfico ou levantamentos geofísicos, são guardados em formato digital e correm risco de obsolescência associado aos suportes onde são armazenados e aos mais diversificados formatos de ficheiros que são utilizados. As práticas associadas ao depósito dos dados resultantes dos trabalhos arqueológicos em repositórios, permitirão a salvaguarda desta informação, única e irrepitível.

No entanto, não basta depositar os dados nos repositórios para que sejam localizáveis, acessíveis e reutilizáveis. Os avanços tecnológicos na captura e armazenamento de dados lançam também novas oportunidades para implementar novas políticas de gestão de dados, de acordo com os princípios FAIR, ou seja, para tornar os dados Localizáveis, Acessíveis, Interoperáveis e Reutilizáveis. Assim, a era dos dados em formato digital proporciona novas oportunidades, mas também novos desafios. O projeto internacional SEADDA, onde Portugal é parceiro e que conta com a participação de 32 países, tem como principal objetivo desenvolver entendimentos comuns sobre as melhores práticas internacionais para a preservação, disseminação e reutilização de dados arqueológicos. A implementação de boas práticas de gestão de dados, associada ao depósito dos dados arqueológicos em repositórios, permite a salvaguarda de informação única e irrepitível e possibilita a disponibilização dos

dados em acesso aberto e a sua interoperabilidade com estruturas internacionais. Estas coleções de dados podem também atuar como catalisadores para envolver diferentes comunidades de conhecimento e interesses, incentivando uma arqueologia mais colaborativa (Wright, Richards, 2018).

Como resultado da investigação que a Unidade de Arqueologia da Universidade do Minho (UAUM) tem desenvolvido ao longo dos últimos anos e com a aprendizagem resultante do envolvimento em projetos internacionais como o SEADDA e ARIADNE, apresentamos neste trabalho a metodologia utilizada para disponibilizar num repositório, um conjunto de dados sobre as moedas resultantes do espólio de escavações arqueológicas realizadas em Braga sob a responsabilidade da UAUM. Começaremos por expor a metodologia utilizada para o registo no sistema de informação (2ArchIS) dos dados arqueológicos, desde os registos de campo à informação complementar processada em gabinete, passando pela análise dos dados e a sua publicação. Finalizamos com a demonstração do processo de exportação de dados do 2ArchIS para o DataRepositóriUM, plataforma onde a informação é disponibilizada a toda a comunidade, contribuindo desta forma para promover a ciência aberta e a reutilização de dados.

2. REGISTO DE DADOS NO SISTEMA DE INFORMAÇÃO 2ARCHIS

No âmbito do Projeto de Salvamento de *Bracara Augusta*, a UAUM iniciou, em 1996, o processo de informatização dos registos provenientes das escavações arqueológicas. O primeiro sistema de registo da informação alfanumérica foi desenvolvido com ferramentas *Microsoft Access*, tendo migrado para uma base de dados relacional em finais de 1999. O sistema criado, denominado SIABRA (Martins, Giestal, 2000), viria a dar origem ao sistema de informação 2ArchIS, desenhado para permitir a integração da informação alfanumérica dos registos de campo com imagens e desenhos vetoriais dos registos gráficos das escavações arqueológicas. O resultado do tratamento desta informação, como as plantas representativas de diferentes fases de ocupação, alçados dos edifícios estudados e reconstituições 3D, passaram também a ser integrados no sistema (Botica, Martins, 2008). Estes conjuntos de dados, quando depositados num repositório, para que sejam acessíveis e reutilizáveis devem obedecer a princípios de dados FAIR (*Findable, Accessible, Interoperable e Reusable*) e devidamente associadas aos contextos em que se inserem. Desta forma, aumentará a confiança para novas reutilizações dos dados e será ofertada maior visibilidade aos seus autores e aos trabalhos realizados.

2.1. Estrutura de dados

O processo de criação de dados digitais assenta no sistema de informação 2ArchIS, que foi desenhado como um sistema modular que interliga informação sobre o território, os sítios arqueológicos ou arquitetónicos, os dados das intervenções realizadas e toda a documentação associada (Figura 1).

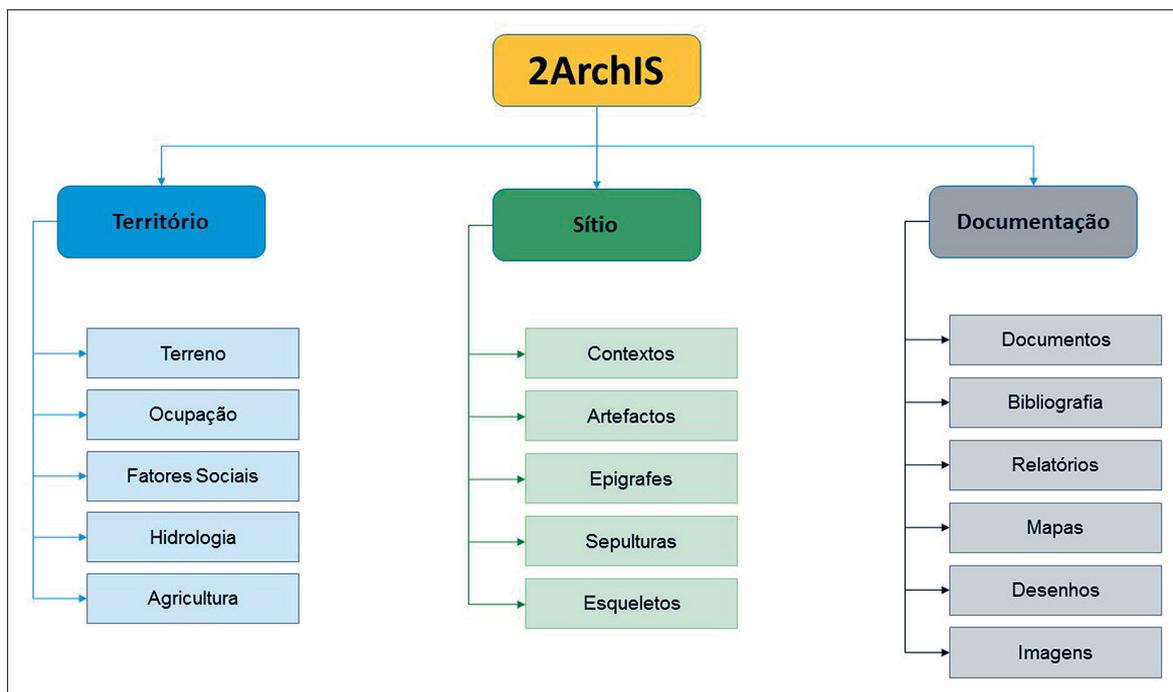


Figura 1

Organograma das entidades do sistema de informação 2ArchIS

A estrutura de dados do sistema 2ArchIS interliga todos os itens de informação recolhidos nas várias etapas do processo de investigação em Arqueologia, suportada por uma base de dados relacional MySQL. Assim, qualquer elemento do 2ArchIS, como, por exemplo um achado de campo do tipo «moeda», tem ligações estabelecidas com o contexto geográfico e sítio arqueológico, bem como aos dados da sua caracterização e da documentação recolhida e produzida durante e após os trabalhos arqueológicos, de acordo com o modelo de dados apresentado na Figura 2.

O registo de dados no 2ArchIS é feito utilizando uma aplicação de *back office*, que tem na origem uma base de dados relacional MySQL, desenvolvida em PHP e HTML, com uma interface *web*, que permite a sua utilização em gabinete, mas também em campo, com recurso a dispositivos móveis (Botica, Martins, 2008). A opção pelo desenvolvimento na UAUM de uma base de dados e da aplicação de registo e gestão da informação foi determinada pelo princípio de utilização de ferramentas *open source* e de interfaces com o utilizador assentes numa plataforma web e independentes dos dispositivos e do local de acesso. Assim, a aplicação de *back office* desenvolvida está acessível num *browser* já instalado nos equipamentos dos utilizadores (Explorer ou Chrome), não requerendo qualquer instalação de *software* específico nem atualizações de versões, quer no campo quer em gabinete. A estrutura desenhada e o desenvolvimento feito internamente também nos garantem autonomia e rapidez na implementação de novos módulos ou para proceder a alterações, ficando imediatamente disponíveis para todos os utilizadores. Por outro lado, a interligação e a modularidade do sistema oferece enormes possibilidades de interligação com outras

bases de dados de projetos de interpretação arqueológica, de estudo das cidades (Botica, 2017; Magalhães, 2019) ou territórios (Maciel *et al.*, no prelo) e com estruturas de dados internacionais, nomeadamente o ARIADNE.

Apresentamos a seguir o processo de registo de dados arqueológicos, para o caso das moedas, através dos formulários de registo do *back office*, recolhidos durante as escavações arqueológicas e, em gabinete, após o processo de limpeza, tratamento e estudo das moedas.

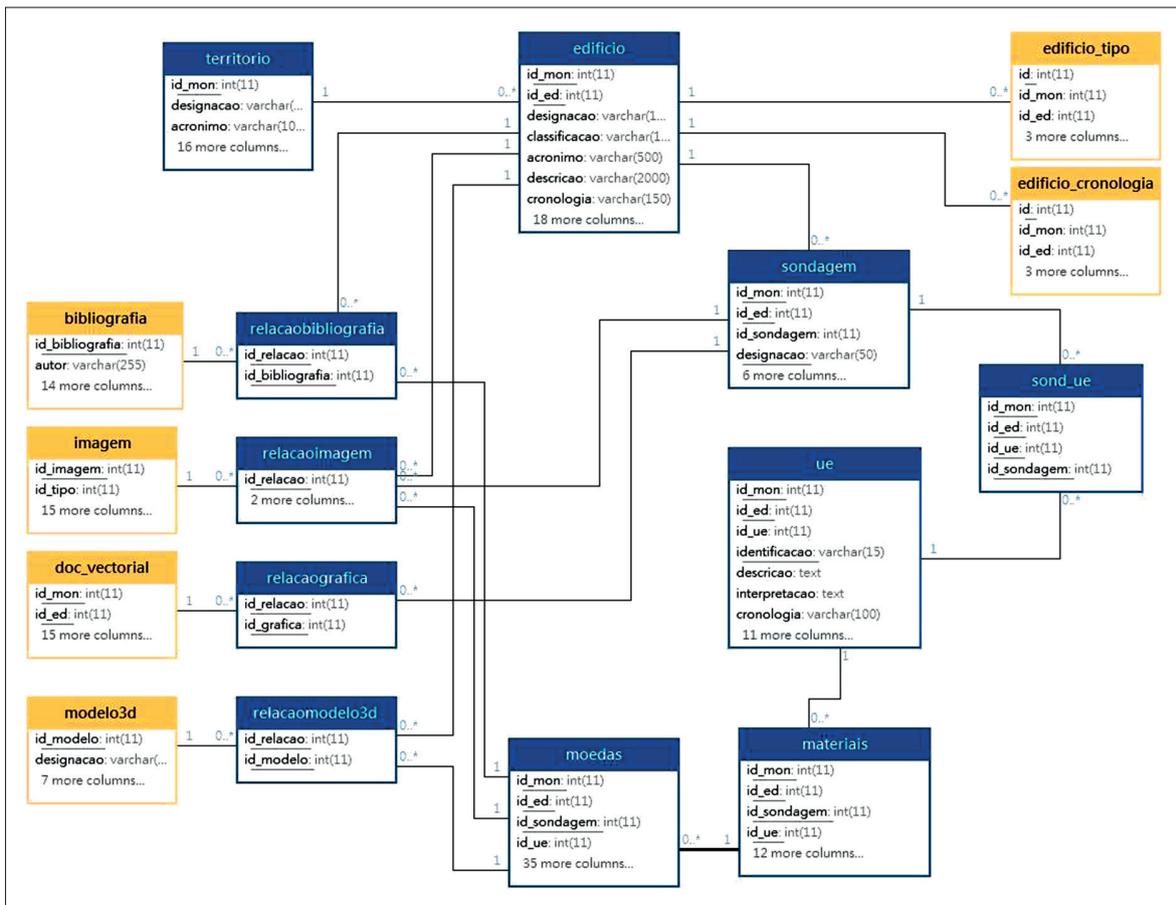


Figura 2

Diagrama UML do modelo de dados do 2ArchIS para o registo de moedas

2.2. Registo de dados

A metodologia utilizada para o registo de achados contempla o registo de dados de campo e a incorporação de elementos resultantes de estudos detalhados dos materiais, passando pela apresentação de possibilidades de análises dessa informação.

Uma vez que os materiais são recolhidos em contexto de escavação arqueológica, a sua introdução no sistema 2ArchIS inscreve-se no processo de registo de todos os dados do sítio arqueológico, como sondagens, unidades estratigráficas e artefactos.

A este registo dos elementos de campo, correspondente à fase inicial de criação de dados, associam-se também fotografias, cartografia, documentação textual e bibliografia (Figura 2). Já em gabinete, após a realização dos trabalhos de campo, inicia-se o processamento dos dados, que incluem desenhos vetoriais de plantas, alçados e perfis, também incorporados ao sistema. A análise dos dados de campo e gabinete produz um conjunto de informação que vai desde o estudo e caracterização dos artefactos à datação de estruturas e fases construtivas bem como a elaboração de plantas interpretadas.

The screenshot shows the '2ArchIS - Unidade de Arqueologia' back office interface. At the top, there is a navigation menu with options: Monumento, Intervenção, Sondagem, UE, Matriz UE, Materiais, Sepulturas, Esqueletos, Epigrafes, and Gliptografia. The main form is divided into three color-coded sections:

- Context (Blue):** Includes 'Sondagem' (S6) and 'UE' (151).
- Field Registration (Green):** Includes 'Nº inventário' (011.005), 'Nº achado' (011.005), 'Descrição' (Ae3 de 357-363), 'Coordenada X' (-24439.800000), 'Coordenada Y' (208833.910000), 'Cota' (192.610000), 'Depósito Legal', 'Nº de moedas' (1), 'Material' (Bronze), 'Valor nominal' (Ae3), 'Eixo da Moeda' (Horizontal), and 'Conservação'.
- Specific Coin Details (Grey):** Includes 'Tipo anverso' (Busto laureado com armadura à direita), 'Tipo reverso' (Imperador em traje militar com globo e lança voltado à esquerda), 'Legenda anverso' (Ilegível), 'Legenda reverso' ((spes rei-publice)), 'CECA' (Alexandria), 'Legenda CECA' ((ale-)), 'Entidade emissora' (Constantius II ou Iulianus II), 'Série' (Imperador sozinho), 'Tesouro' (011), and 'Imitação'.

Additional sections on the right include 'Cronologia' (Período: Romano - Baixo Imperial, Século: IV, Datação: 357-363), 'Dimensões' (Espessura: 1 mm, Peso: 1.5 g, Diâmetro: 16 mm), and 'Tratamento' (Limpeza: Com álcool puro, Conservação: Benzotriazol 3%, Consolidação: Paraloid B72 3% c/ xileno (Ac), Restauro).

At the bottom, there are buttons for 'Gravar e Voltar' and 'Cancelar', and a note: '* Campos obrigatórios; NOTA: nas coordenadas usar . para casas decimais'.

Figura 3
Formulário do *back office* para o registo de dados de moedas

O registo de achados arqueológicos na aplicação de *back office* é feito selecionando uma entrada específica por tipo de objeto. Embora todos os materiais sejam registados em contexto de escavação arqueológica com atributos comuns, como a sondagem, unidade estratigráfica (UE) e tipo, posteriormente, estes são caracterizados em gabinete com dados específicos de acordo com a sua tipologia. O formulário apresentado na figura 3 é referente ao registo de moedas, organizado em três grupos distintos de dados. O primeiro grupo, destacado a azul, contém os campos que associam a moeda ao contexto da escavação arqueológica do sítio, sondagem e unidade estratigráfica (UE), cuja caracterização já foi feita em formulários próprios. O segundo grupo, destacado a verde, é composto por informação de registo de campo, comum a

todos os artefactos, independentemente do seu tipo (moeda, cerâmica, vidro, etc.), designadamente uma descrição breve com o nº de campo do(s) achado(s), sempre que se justifica as coordenadas geográficas da sua localização e o depósito legal. O terceiro grupo, correspondente a todos os restantes campos, permite registar os dados específicos para o tipo de artefacto, neste exemplo para as moedas, produzidos na sequência das análises realizadas após a limpeza e/ou tratamento dos materiais.

Os campos de registo foram definidos seguindo as normas comumente utilizadas pela comunidade científica, distinguindo-se campos de texto livre e campos associados a uma lista de valores (LOV), onde o utilizador apenas pode escolher uma opção disponibilizada na lista. É o caso de atributos como o tipo de material, a oficina (CECA), a entidade emissora, a série, o período cronológico e século. São atributos onde as listas de valores ajudam na decisão de escolha dos descritores e nos garantem que o vocabulário usado segue os padrões adotados, evitando erros de digitação ou utilização de vocábulos diferentes para designar o mesmo atributo.

Já os campos como tipo e legenda, anverso e reverso e legenda CECA são de texto livre, uma vez que se tratam de campos onde o utilizador regista aquilo que vê. Estes atributos descritivos identificam os ícones e símbolos presentes nos aversos e reversos das moedas, tendo-se adotado a metodologia de registar em letras maiúsculas a transcrição das legendas visíveis e, sempre que possível, em letras minúsculas e entre parênteses a restituição das letras ilegíveis. No entanto, são estes campos de texto livre que possibilitam a identificação e o registo de campos normalizados e associados a listas de valores.

Relativamente às cronologias optámos por usar três campos que correspondem ao período cronológico e subdivisão em século e datação. Nos dois primeiros, período (Romano - Baixo Imperial) e século (IV), decidimos por associar listas de valores predefinidas, enquanto que na datação as possibilidades são infinitas, tendo sido feita a opção de registo em texto livre, colocando-se a data num intervalo possível, com a data de início e fim separadas por um hífen (357-363).

Em relação aos dados dos atributos arqueométricos das moedas, estes reportam ao valor máximo observado no numisma e foram registados para as categorias espessura e diâmetro, medidas em milímetros, e peso, expresso em gramas.

A informação acerca dos produtos utilizados para tratamento do material é registada nos campos associados à limpeza, conservação e consolidação do numisma.

A cada moeda pode ainda ser associada bibliografia e imagens do anverso e reverso.

Esta forma de registo permite uma grande flexibilidade para estabelecer critérios de pesquisa para elaboração listagens, estatísticas e ligação a Sistemas de Informação Geográfica (SIG), como a seguir se exemplifica.

2.3. Acesso e visualização dos dados

A aplicação de *back office* disponibiliza formulários de consulta através de critérios pré-estabelecidos, como por exemplo pesquisa de materiais encontrados em determinada UE, por período cronológico, século ou datação mais fina (anos) ou por outras palavras-chave. Por outro lado, oferece ainda um conjunto de mapas estatísticos da distribuição das moedas por Sondagem/UE, por material ou por entidade emissora. É ainda possível elaborar catálogos, como o apresentado na Figura 4 com o exemplo das moedas do Seminário de S. Pedro e S. Paulo e Museu Pio XII.



Unidade de Arqueologia da Universidade do Minho

Lista de Moedas

Seminário de Santiago

Nº inventário: 011.336

Sondagem: S6

UE: 151

Nº achado: 011.336

Nº objecto:

Depósito:

Nº catálogo: RIC VII Rome 331 (var)

Tesouro: 011

Coordenadas: X: -24439.80

Y: 208833.91

Cota: 192.61

Material: Bronze

Classificação: Moeda romana do Baixo-império

Oficina: Roma (4ª officina)

Ent. Emissora: Serie Urbana




Século: IV

Legenda: RFQ

Cronologia: 330

Valor nominal: Ae3

Série: Comemorativa urbana

	Anverso	Reverso
Tipo	Busto de Roma com elmo plumado, manto imperial e colar à esquerda.	Loba capitolina com Romulo e Remo; duas estrelas.
Legenda	VRBS ROMA	
	Peso: 1.9 Módulo: 18 Grossura: 1.5 Pos. cunhos: Horizontal.	Conservação: Muito Bom
Tratamento :	Limpeza: Com álcool puro	Conservação: Benzotriazol 3%
	Consolidação: Paraloid B72 3% c/ xileno (Acryloid B72)	

Figura 4

Pormenor do catálogo de moedas do Seminário de S. Pedro e S. Paulo e Museu Pio XII

Os dados registados no sistema de informação 2ArchIS são utilizados pela UAUM para produzir relatórios técnicos e publicações científicas, que ficam depositados no repositório institucional da Universidade do Minho (RepositóriUM), entre outras plataformas de divulgação. A grande quantidade de escavações arqueológicas realizadas pela UAUM ao longo de mais 40 anos de atividade resultou num vasto conjunto de dados e de conhecimento colocado ao serviço da investigação em Arqueologia, através do apoio a muitos trabalhos científicos, nomeadamente a dissertações de mestrado e teses de doutoramento.

2.4. Análise dos dados

A análise individual dos materiais configura-se como uma atividade que envolve um enorme dispêndio de energia e tempo, desde a lavagem e marcação, passando pelo registo fotográfico e caracterização. O sistema 2ArchIS tem já registado um considerável conjunto de materiais, desde as cerâmicas, moedas, líticos ou vidros, provenientes de várias intervenções arqueológicas, fomentando possibilidades de análise deveras ampliadas, em particular quando combinadas com ferramentas informáticas que permitem diferentes tipos e níveis de visualização. No âmbito de dissertações de mestrado em Arqueologia na Universidade do Minho, destacamos dois tipos de análises feitas com recurso aos dados inseridos no sistema de informação e que tiveram diferentes perspetivas, culminando em resultados distintos. O primeiro, intitulado «As moedas das Carvalheiras. Contributo para o estudo da circulação monetária», interligou o 2ArchIS com um Sistema de Informação Geográfica (SIG) para mapear o posicionamento dos numismas associados a determinados filtros temporais, e assim visualizar a sua dispersão no interior de uma zona arqueológica. Os mapas produzidos apoiaram o estudo da relação entre a concentração de objetos em determinados setores da casa, a «*domus* das Carvalheiras de Braga», e a evolução arquitetónica do edificado ao longo dos cinco séculos de existência (Amaral, 2007). Num outro trabalho, intitulado «Tesouros numismáticos baixo-imperiais de contextos domésticos de *Bracara Augusta*» (Machado, 2017), estudaram-se os ocultamentos monetários da cidade a partir de diversas categorias, como cronológica ou arqueométrica, mas também geográfica, em que se valorizaram os centros produtores que abasteciam a capital da *Gallaecia* nos séculos III-V. Um dos exemplos de mapas produzidos (Figura 5) evidencia a origem das moedas, produzidas um pouco por todo o Império Romano em cidades como *Constantinopolis*, *Alexandria*, *Roma* e *Arelate*, e o seu impacto na massa monetária resultante de um achado na intervenção do Seminário de S. Pedro e S. Paulo e Museu Pio XII.

Estes dois exemplos, longe de esgotarem as hipóteses de análise do material numismático, apenas assinalam possibilidades de apoio à interpretação com recurso a ferramentas de pesquisa que permitem o cruzamento dos dados das moedas com diferentes variáveis, de uma forma rápida, intuitiva, completa e eficiente.

A gestão, divulgação e partilha de informação tem sido uma constante da UAUM, quer através da ampla divulgação dos seus trabalhos, em relatórios científicos e técnicos, quer pela divulgação de informação através do seu *website* (<http://www.uaum.uminho.pt/webapp/>). No entanto, apesar desta sistemática divulgação, sentimos que chegou o momento de partilhar os nossos conjuntos de dados, como contributo para ciência aberta em Arqueologia. Assim, a UAUM foi uma das primeiras a utilizar o repositório de dados de investigação da Universidade do Minho, o DataRepositóriUM, para salvaguarda e disponibilização de acesso a conjuntos de dados arqueológicos.

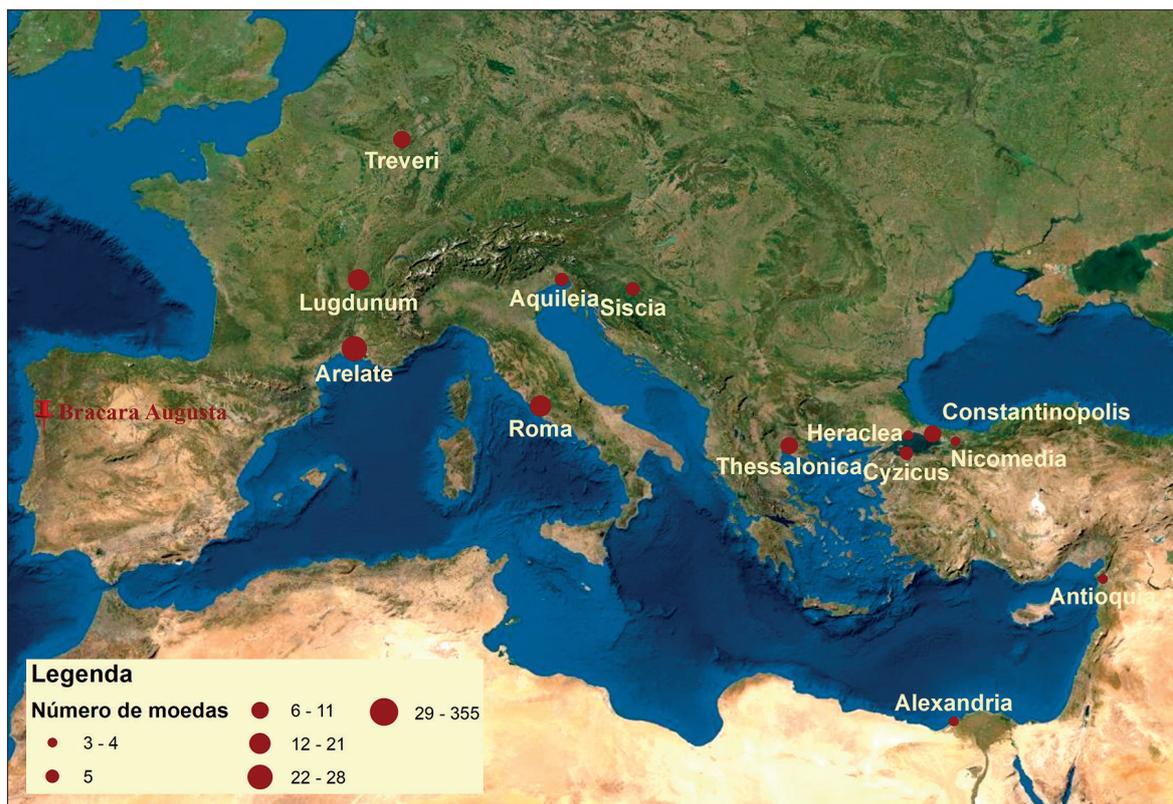


Figura 5

Mapa com a origem das moedas do tesouro do Seminário de S. Pedro e S. Paulo e Museu Pio XII

3. DataRepositóriUM

3.1. Objetivos

A Arqueologia incorpora nos seus procedimentos de trabalho e registo muitos dados que são nativamente digitais. No entanto, os dados das escavações arqueológicas, que resultam de ações irrepetíveis, ou não estão acessíveis ou estão colocados em relatórios e publicações, sendo de difícil localização e reutilização. Para difundir a ciência aberta e a reutilização de dados é necessário que estes estejam disponíveis em repositórios abertos e que garantam a sua acessibilidade e perenidade.

A Universidade do Minho disponibilizou à comunidade académica, em fevereiro de 2020, o repositório de dados da UMinho «DataRepositóriUM», sendo uma plataforma desenvolvida com o *software Dataverse* para partilha, preservação, citação, exploração e análise de dados de investigação (Principe *et al.*, 2018). A Unidade de Arqueologia é responsável pelo depósito de conjuntos de dados de Arqueologia, dos quais destacamos o *Dataverse* «Arqueologia – Moedas» (<https://datarepositorium.uminho.pt/dataverse/arqueologia-moedas>).

Do ponto de vista dos depositantes, o repositório facilita a salvaguarda e a acessibilidade aos dados, bem como a sua reutilização. Para além disso, a associação dos conjuntos de dados a um identificador único e persistente (DOI) e ao(s) seu(s) autor(es) promove ainda as citações e o reconhecimento dos autores e do seu trabalho. Na perspetiva dos utilizadores dos dados, o repositório facilita a reutilização dos dados em novas interpretações ou em novos projetos de investigação (Marwick, Pilaar Birch, 2018).

No entanto, para que os dados depositados no repositório sejam localizáveis, a sua criação e arquivamento deve obedecer a um conjunto de princípios que têm que garantir que os dados sejam FAIR, isto é, Fáceis de encontrar, Acessíveis, Interoperáveis e Reutilizáveis (Hollander *et al.*, 2018).

3.2. Metodologia usada para publicação de dados no DataRepositóriUM

O depósito de dados num repositório tem duas componentes fundamentais: os dados e os metadados. Assim, apresentamos as várias fases do processo de criação de dados e metadados para publicar os conjuntos de dados de moedas, já inseridos no sistema 2ArchIS, de acordo com os princípios FAIR.

As etapas de criação dos ficheiros para os conjuntos de dados das moedas foram as seguintes:

- **Exportação dos dados** da base de dados MySQL do sistema 2ArchIS. O ficheiro a publicar no DataRepositóriUM resultou de uma exportação para um ficheiro do tipo «csv» dos registos inseridos na base de dados. Este formato de arquivo, para além de ser recomendado pelo repositório, é um formato sustentável, largamente usado e entendido pela comunidade. O nome atribuído ao ficheiro «ListaMoedas» é acompanhado por um breve descritor que indica com clareza o tipo e quantidade de dados, «Catálogo das 481 moedas do tesouro numismático do Seminário de S. Pedro e S. Paulo e Museu Pio XII».

Neste ficheiro «ListaMoedas.csv» incluímos os nomes dos atributos em português e em inglês, com todos os dados registados em campo e em gabinete. A normalização utilizada no registo dos dados, através de formulários uniformizados e com vocábulos comumente aceites e utilizados pela comunidade, para além de dar consistência aos dados, possibilita que o repositório seja interoperável com outros conjuntos de dados. Cada linha do ficheiro csv, para além dos dados de caracterização da moeda, contém ainda dados do seu contexto, como o sítio arqueológico, sondagem, unidade estratigráfica e coordenadas geográficas da sua localização (Figura 6).

Incluímos ainda neste conjunto de dados, que foi exportado diretamente da base de dados para um ficheiro do tipo pdf, o catálogo das moedas, já apresentado na Figura 4, com o objetivo de estabelecer a correspondência entre os dados e as fotos de anverso e reverso das moedas.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
Site	Survey	Stratum	Inventory num	Century	Chronology	Denominat	Person	Material	Mint	Mintmark	Anverse	Anverse lege	Obverse	Obverse l	Weight
Sítio	Sondagem	UE	nº inventario	Século	Cronologi	Classificaç	Entidade emissora	Material	ceca	Legenda ceca	Anverso	Legenda anve	Reverso	Legenda r	Peso
Seminário S. Pedro, S 56			151 11.001	IV	353-358	Moeda roi Constantius Gallus ou Consta		Bronze	Indeterminada	Illegível.	Busto laureado com Illegível.	Soldado a at. (fel temp			1,9
Seminário S. Pedro, S 56			151 11.002	III-V		Moeda roi Indeterminada		Bronze	Indeterminada	Illegível.	Busto laureado com Illegível.	Illegível.	Illegível.		1,3
Seminário S. Pedro, S 56			151 11.003	III-V		Moeda roi Indeterminada		Bronze	Indeterminada	Illegível.	Busto laureado com Illegível.	Illegível.	Illegível.		0,9
Seminário S. Pedro, S 56			151 11.004	IV	367-375	Moeda roi Constans		Bronze	Constantinopolis	CON(-)	Busto de Valens lau (dn va)LENS P Imperador a GLORIA (n				1,6
Seminário S. Pedro, S 56			151 11.005	IV	357-363	Moeda roi Constantius II ou Iulianus II		Bronze	Alexandria	(ale-)	Busto laureado com Illegível.	Imperador e (spes rei-			1,5
Seminário S. Pedro, S 56			151 11.006	IV	353-358	Moeda roi Constantius II ou Iulianus II		Bronze	Indeterminada	Illegível.	Busto laureado à d(-)	Soldado a at. (fel temp			1,9
Seminário S. Pedro, S 56			151 11.007	IV	353-354	Moeda roi Constantius II		Bronze	Indeterminada	Illegível.	Busto de Constantir DN CONSTAN Soldado a at. (fel t)EMP				1,6
Seminário S. Pedro, S 56			151 11.008	IV	378-383	Moeda roi Gratianus		Bronze	Arelate (3ª officin	TCON	Busto de Gratianus (D N GRATIA-(Coroa de lou VOT XV M				1,4
Seminário S. Pedro, S 56			151 11.009	IV	364-378	Moeda roi Valens		Bronze	Indeterminada	Illegível.	Busto de Valens dia (d n val)EN-S. Victoria a avr (securitas				1,2
Seminário S. Pedro, S 56			151 11.010	IV	342-348	Moeda roi Constantius II		Bronze	Arelate (1ª officin	PAR(I)	Busto de Constantii C(constanti-us Duas Victori VICTORIAI				1,2
Seminário S. Pedro, S 56			151 11.011	V	402-423	Moeda roi Honorius		Bronze	Aquileia	AO(-)	Busto de Honorius DN HONORI(- Victoria avar (salus rel-				0,8
Seminário S. Pedro, S 56			151 11.012	IV	342-348	Moeda roi Constantius II		Bronze	Aquileia ou Roma	Illegível.	Busto de Constantii (constanti-IV: Duas Victori (victori)AI				1

Figura 6

Exemplo do ficheiro csv das moedas do tesouro do Seminário de S. Pedro e S. Paulo e Museu Pio XII

- **Metadados.** Os metadados são essenciais para tornar os dados localizáveis, em particular os que são usados para citar e descrever dados. Os Metadados registados incluem uma descrição do conjunto de dados, o título e subtítulo, autor(es), palavras chave, tipo de dados, período a que reportam os dados e ainda *links* para documentação e publicações associadas. Ao depositar os dados é automaticamente atribuído um identificador persistente a este conjunto de dados (DOI) e associadas datas de produção dos dados, de depósito e de publicação.

A Figura 7 apresenta os ficheiros disponibilizados no DataRepositóriUM para os conjuntos de dados das moedas do Seminário de S. Pedro e S. Paulo e Museu Pio XII. Para além dos ficheiros já descritos: catalogo.pdf e ListaMoedas.csv, junta-se ainda um ficheiro README.txt que fornece informação adicional para que os dados possam ser acedidos e compreendidos ao longo do tempo.

The screenshot shows the DataRepositóriUM interface with a search filter set to 'File Type: All' and 'Access: All'. There are 3 files listed:

- Catalogo.pdf**: Adobe PDF - 17.4 MB - Feb 27, 2020 - 7 Downloads. MD5: 3fad12967f83a4a5e96fe348ebdd80. Description: Catálogo das 481 moedas do tesouro numismático do Seminário S. Pedro, S. Paulo e Museu PIO XII.
- ListaMoedas.csv**: Comma Separated Values - 138.5 KB - Mar 10, 2020 - 5 Downloads. MD5: 93df7f75d62169561e3276d696602177. Description: Dados das 481 moedas do tesouro numismático do Seminário S. Pedro, S. Paulo e Museu PIO XII.
- README.txt**: Plain Text - 1.2 KB - Feb 27, 2020 - 5 Downloads. MD5: a2ebfb7d461801dff0d4c245a9589f7d9.

Figura 7

Ficheiros do DataRepositóriUM para as moedas do Seminário de S. Pedro e S. Paulo e Museu Pio XII

4. CONCLUSÃO

O registo e partilha de dados, para além da sua função de gestão e salvaguarda, torna possível a sua reutilização nas mais variadas análises, sem que cada nova utilização dos dados implique tempo consumido na sua digitalização. No entanto, os conjuntos de dados a partilhar, para poderem ser reutilizados, deverão obedecer aos princípios FAIR, promovendo a consistência dos dados, a sua localização e interoperacionalidade com outros dados. Desta forma, acreditamos que aumentarão as possibilidades de uso e de reflexão sobre os dados, terreno fértil para o progresso de investigações e o avanço científico do conhecimento em Arqueologia.

Da experiência de depósito de publicação de dados de achados arqueológicos, para o caso apresentado das moedas das escavações arqueológicas da UAUM, podemos concluir que o trabalho prévio realizado durante o seu registo, nomeadamente usando uma estrutura de dados relacional e com campos e vocabulários normalizados, foi fundamental. Por um lado, simplifica o processo de geração dos ficheiros de arquivos e, por outro, permite que estes dados possam ser entendidos pelos futuros utilizadores, quer nas suas características e propriedades, quer na sua contextualização temporal e geográfica. A associação dos conjuntos de dados a *links* para documentação e publicações associadas permite que futuras reutilizações possam ter informação adicional sobre resultados de análises e trabalhos realizados com os dados, quer para validação quer para produção de novos estudos ou novas interpretações.

As vantagens expostas para a criação e utilização dos repositórios de dados são vastas, uma vez que estes arquivos digitais podem ser partilhados e acedidos muitas vezes, em qualquer local e ao longo do tempo, facilitando novos trabalhos e colaborações partilhadas. No entanto, o sucesso da sua utilização em larga escala passará muito pela adesão dos autores dos dados, pela implementação de interfaces amigáveis e pela confiança que esses dados oferecem aos futuros utilizadores.

REFERÊNCIAS BIBLIOGRÁFICAS

- AMARAL, L. (2007). *As moedas das Carvalheiras. Contributo para o estudo da circulação monetária em Bracara Augusta*, UAUM/Narq, Braga.
- BOTICA, N. (2017). «Contributo do sistema de informação 2ArchIS para o conhecimento das cidades romanas de Braga e Lugo» *Philtáte* Volume 2, 387-397, Lugo.
- BOTICA, N., MARTINS, M. (2008). «Sistemas de informação em arqueologia - a experiência de Bracara Augusta», *Férvendes* nº 5, 11-14, Vilalba.
- HOLLANDER, H., MORSELLI, F., ADMIRAAL, F., UITERWAAL, F., NOORDEGRAAF, M. (2018). *Guidelines to FAIRify data management and make data reusable: PARTHENOS*.
- LAUŽIKAS, R., DALLAS, C., THOMAS, S., KELPŠIENĖ, I., HUVILA, I., LUENGO, P., NOBRE, H., TOUMPOURI, M., VAITKEVIČIUS, V. (2018). «Archaeological Knowledge Production and Global Communities: Boundaries and Structure of the Field», *Open Archaeology* Volume 4, 350–364.
- LODWICK, L. (2019). «Sowing the Seeds of Future Research: Data Sharing, Citation and Reuse in Archaeobotany» *Open Quaternary* 5, p. 7, 1-15.
- MACHADO, D. (2017). *Tesouros numismáticos baixo-imperiais de contextos domésticos de Bracara Augusta*, Universidade do Minho, Braga.
- MACIEL, S., BOTICA, N., BLANCO ROTEÁ, R. (no prelo). «A Paisagem Romanizada da Citânia de Sanfins, Paços de Ferreira (Portugal). Proposta metodológica para uma análise macro espacial», *Glyphos* 7, Zamora.
- MAGALHÃES, M. (2019). *A domus Romana no Noroeste Peninsular: Construção, Arquitetura e Sociabilidades*, Tese de Doutoramento, Universidade do Minho, Braga.
- MARTINS, M. GIESTAL, C. (2000). «O Projecto SIABRA: um Sistema de Informação Geográfica para a Arqueologia Urbana em Braga», *3º Congresso de Arqueologia Peninsular. SIG's aplicados à Arqueologia da Península Ibérica* (Vila Real, 1999), 43-61, Porto.
- MARWICK, B., PILAAR BIRCH, S. (2018). «A Standard for the Scholarly Citation of Archaeological Data as an Incentive to Data Sharing», *Advances in Archaeological Practice* 6(2), 125-143, Cambridge.
- PRINCIPE, P., GOMES, P., RODRIGUES, E. (2018). «DataRepositóriUM: projeto de implementação do repositório de dados para a Universidade do Minho», *4º Fórum - Gestão de dados de investigação* (Castelo Branco, 2018). Obtido em 14 de 05 de 2020, de https://forumgdi.rcaap.pt/wp-content/uploads/2018/09/02_4ForumGDI_datarepositorium_final.pdf.
- ŠTULAR, B., CORNS, A. (2014). «Impressions from the ARIADNE Community», *Italian Semester of Presidency of the European Union International Conference on Research Infrastructures and e-Infrastructures for Cultural Heritage* (Roma, 2014).
- WRIGHT, H., RICHARDS, J. (2018). «Reflections on Collaborative Archaeology and Large-Scale Online Research Infrastructures», *Journal of Field Archaeology* Volume 43, 60-67, Londres.

