



Pecha Kucha

DOI: [10.21680/2447-7842.2023v9n2ID33822](https://doi.org/10.21680/2447-7842.2023v9n2ID33822)

Certificação CoreTrustSeal: a qualificação dos processos do DataRepositóriUM rumo a dados FAIR e a curadores de dados mais capacitados

CoreTrustSeal certification: the qualification of DataRepositóriUM's processes towards FAIR data and more capable data curators

Pedro Príncipe ¹

André Vieira ²

Yulia Karimova ³

Paula Moura ³

| | | |
|--------------------------|--------------------------------|--------------------------|
| Submetido em: 17/04/2023 | Aprovado na ConfOA: 14/06/2023 | Publicado em: 04/12/2023 |
|--------------------------|--------------------------------|--------------------------|

Resumo: O Repositório Institucional de dados de investigação da Universidade do Minho – DataRepositóriUM – iniciou o processo de preparação para obtenção da certificação CoreTrustSeal em 2022 e procedeu à submissão da avaliação dos requisitos em 2023. Este trabalho detalha a metodologia seguida no processo de preparação para cumprimento dos requisitos formais do CoreTrustSeal, as lições aprendidas com este processo de qualificação do DataRepositóriUM e os principais resultados obtidos. A certificação de confiabilidade, estabelecida pelo CoreTrustSeal, foi um dos objetivos definidos para os primeiros anos de operação do serviço, tendo a equipa responsável dos Serviços de Documentação e Bibliotecas da Uminho (SDB) desenvolvido um processo de preparação da certificação e de auto-avaliação que incluiu: i) Auto-diagnóstico e identificação das lacunas; ii) Documentação dos processos e políticas associados aos requisitos; iii) Criação dos processos e

¹ Universidade do Minho.

² Universidade do Minho.

³ Universidade do Minho.



políticas de serviço associados aos requisitos por cumprir ou pouco definidos; iv) Disponibilização e agregação da informação e documentação dos processos na página de suporte ao DataRepositóriUM; v) Associação dos principais processos, atividades e políticas definidas no quadro da certificação CoreTrustSeal ao Sistema de Gestão de Qualidade da USDB. Com este artigo espera-se dar uma visão holística do processo de certificação CoreTrustSeal associado a um repositório.

Palavras-chave: certificação de repositórios; CoreTrustSeal; DataRepositóriUM; dados FAIR; curadoria de dados.

Abstract: The University of Minho's Institutional Research Data Repository - DataRepositóriUM - started the preparation process to obtain CoreTrustSeal certification in 2022 and submitted the requirements assessment in 2023. This presentation details the methodology followed in the preparation process to meet the formal requirements of CoreTrustSeal, the lessons learned from this qualification process for DataRepositóriUM and the main results obtained. Reliability certification, established by CoreTrustSeal, was one of the objectives set for the service's first years of operation, with the team responsible from the UMinho's Documentation and Library Services developed a process to prepare for certification and self-assessment, including: i) Self-diagnosis and gaps identification; ii) Documentation of the processes and policies associated with the requirements; iii) Creation of the service processes and policies; iv) Publication of the support page with information and documentation of the DataRepositóriUM processes; v) Association of the main processes, activities and policies within the framework of the CoreTrustSeal certification with the USDB Quality Management System. The aim of this article is to provide a holistic view of the CoreTrustSeal certification process associated with a institutional repository.

Keywords: repository certification; CoreTrustSeal; DataRepositóriUM; FAIR data; data curation.



1 INTRODUÇÃO

O Repositório Institucional de dados de investigação da Universidade do Minho – DataRepositóriUM – iniciou o processo de preparação para obtenção da certificação CoreTrustSeal (2022) e procedeu à submissão da avaliação dos requisitos em 2023. Esta apresentação detalha a metodologia seguida no processo de preparação para cumprimento dos requisitos formais do CoreTrustSeal, as lições aprendidas com este processo de qualificação do DataRepositóriUM e os principais resultados obtidos (Universidade do Minho, 2022).

O mote para este trabalho está alicerçado no pressuposto de que a obtenção da certificação CoreTrustSeal, enquanto instrumento de auto-avaliação e validação externa, vale mais pelo processo iterativo de melhoria dos procedimentos de trabalho, da documentação associada e da definição de políticas e condições de prestação do serviço.

O CoreTrustSeal é a certificação mais utilizada para ajudar a estabelecer repositórios de dados fidedignos. O processo de candidatura para obter a certificação inclui uma auto-avaliação por pares das instalações dos repositórios digitais, da organização dos procedimentos e das políticas do serviço de repositório.

O DataRepositóriUM, disponibilizado à comunidade da Universidade do Minho em fevereiro de 2020, é o repositório de dados institucional da UMinho que serve para partilhar, publicar e gerir dados de investigação gerados e recolhidos pela atividade dos investigadores e nas unidades de investigação da Universidade do Minho. Apresenta-se com os seguintes objetivos: i) Depositar ficheiros em datasets e coleções organizadas; ii) Preservar ficheiros e datasets; iii) Apoiar o cumprimento dos requisitos de financiadores; iv) Registrar identificadores persistentes (DOI); v) Permitir o acesso controlado e sob identificação prévia; vi) Associar licenças de reuso e explicitar os termos de uso; vii) Documentar os datasets com metadados disciplinares; viii) Facilitar a disseminação dos resultados de investigação; ix) Viabilizar a citação de dados e o crédito aos produtores; x) Facilitar a disponibilização de dados aos editores⁴.

⁴ Missão e objetivos do DataRepositóriUM: <https://datarepositorium.uminho.pt/about/missao.html>



A melhoria contínua da prestação do serviço, quer no domínio da maior visibilidade e utilização pelos potenciais utilizadores da comunidade da UMinho, quer no domínio dos procedimentos, fluxos de trabalho e documentação, foi um objetivo assumido pela coordenação do serviço nas Bibliotecas da UMinho. A certificação de confiabilidade, estabelecida pelo CoreTrustSeal, foi por isso mesmo um dos objetivos definidos para os primeiros anos de operação do serviço, a par do registo do serviço nos principais diretórios mundiais e sistemas de agregação de metadados, como o RE3DATA, FAIRsharing ou OpenAIRE (2020).

Nas Diretrizes sobre Gestão de Dados FAIR nos programas de financiamento da ciência Horizonte 2020 e no Horizonte Europa, a Comissão Europeia define que os dados e metadados associados, a documentação e software, devem ser depositados preferencialmente em repositórios certificados que suportem o acesso aberto. Deste modo, os investigadores e gestores de informação podem confiar num quadro de várias normas internacionais de certificação para repositórios digitais. No quadro europeu de auditoria e certificação de repositórios digitais, estão disponíveis três instrumentos de certificação, com graus de complexidade e profundidade crescentes: a) CoreTrustSeal (CTS): baseado no Data Seal of Approval (DSA) e World Data System (WDS); b) Selo Nestor: verificação de acordo com DIN 31644; c) Certificação ISO 16363.

As avaliações variam em intensidade, desde uma revisão por pares da documentação (auto-avaliação), até uma visita preparada no local por uma equipa de auditoria externa. Estes instrumentos são utilizados em todo o mundo. Existem também repositórios de dados com uma base de utilizadores sólida e duradoura, no entanto, é esperado que estes repositórios se candidatem à certificação num futuro próximo, porque a organização do financiamento da ciência e as organizações que realizam investigação na Europa (financiadores coordenados na Science Europe) estão a desenvolver critérios para a seleção de repositórios de confiança.



2 METODOLOGIA

A certificação do CoreTrustSeal é um processo agregador de auto-avaliação e validação que integra a certificação mais ampla do Nestor-Seal DIN 31644 e a certificação mais formal da ISSO 16363. O CoreTrustSeal resulta de esforços globais da comunidade de repositórios e arquivos de dados para assegurar um quadro de referência global de certificação de repositórios que inclui as certificações existentes, de nível mais alargado do Nestor-Seal e de nível formal da ISO. Para além deste foco no serviço de repositório digital de dados, o CoreTrustSeal pode fornecer certificação de nível central para outras entidades de investigação, tais como serviços de dados e software⁵.

A equipa do gabinete de Gestão de Informação Científica, Repositórios e Ciência Aberta dos serviços de Documentação e Bibliotecas da UMinho (USDB, 2023) desenvolveu um processo cuidadoso de preparação da certificação e de auto-avaliação. Este processo teve início em abril de 2022, a partir de uma atividade conjunta com os responsáveis do Repositório de Dados do Instituto Politécnico de Bragança, e culminou com a submissão para a solicitação de Certificação de Repositório de Dados CoreTrustSeal em abril de 2023. A equipa da UMinho procedeu da seguinte forma:

- 1) Estudo e documentação da visão global dos requisitos CoreTrustSeal;
- 2) Auto-diagnóstico e identificação das lacunas mais críticas para o cumprimento dos requisitos;
- 3) Documentação dos processos e políticas associados aos requisitos parcialmente cumpridos, mas não devidamente documentados;
- 4) Criação dos processos e políticas de serviço associados aos requisitos por cumprir ou pouco definidos;
- 5) Disponibilização e agregação da informação e documentação dos processos numa página web pública de suporte ao DataRepositóriUM;

⁵ Referência referida no site de apresentação do CTS: <https://www.coretrustseal.org/about/>



- 6) Associação dos principais processos, atividades e políticas definidas no quadro da certificação CoreTrustSeal ao Sistema de Gestão de Qualidade da USDB.

A atividade de qualificação como repositório confiável está enquadrada no projeto PortAberta – Portal de Investigação para a Ciência Aberta – uma operação no âmbito do Sistema de Apoio à Modernização e Capacitação da Administração Pública (SAMA2020) da Agência para a Modernização Administrativa (AMA). Neste projeto, há uma área de trabalho em desenvolvimento que se centra nos serviços de apoio à gestão de dados de investigação e a integração de registos de datasets e planos de dados como um produto relevante para o Portal de Investigação. As tarefas desta área de trabalho visam apoiar a estratégia para a gestão de dados nas duas Instituições, com particular enfoque em providenciar uma maior centralidade aos Repositórios, onde se inclui os repositórios de dados do software Dataverse, que ambas as instituições já possuem, e para o qual era prioritário qualificar com melhores processos de curadoria e preservação, procurando obter certificação de confiabilidade alinhada com as melhores práticas internacionais.

3 RESULTADOS

O primeiro resultado do processo de auto-avaliação para a solicitação de certificação CoreTrustSeal está relacionado com o potencial formativo do processo em si mesmo. Neste sistema iterativo levado a cabo entre as duas instituições e no quadro de um trabalho intenso e colaborativo na equipa da USDB da UMinho, fica fortemente vincado o papel pedagógico do instrumento de auto-avaliação que a certificação CoreTrustSeal configura. De facto, o processo de auto-diagnóstico e de avaliação de cada um dos requisitos estabelecidos na norma funciona como mecanismo auto-formativo, e quando assumido de forma colaborativa, como foi na pequena equipa da USDB, exponencia o potencial de capacitação dos profissionais a exercerem funções de curadoria de (meta)dados no DataRepositóriUM.



Em segundo lugar, valerá a pena destacar e explorar a importância, para o cumprimento de requisitos da certificação, da integração de aspetos do sistema de informação organizacional e de políticas mais transversais da Universidade. O instrumento de auto-avaliação obriga a considerar a inclusão de componentes, procedimentos e políticas de outras unidades dentro da UMinho. Desde logo o alinhamento com os procedimentos de proteção de dados na instituição, particularmente relacionados com o tratamento dos dados pessoais no cumprimento das regras para proteção das pessoas singulares previstas na legislação. Para além deste aspeto, a necessidade de alinhar de forma explícita com o Código de Conduta Ética da UMinho, salvaguardando os princípios éticos que enquadram as atividades de ensino, de investigação e de interação com a sociedade também nos processos e políticas do serviço DataRepositóriUM.

Por último, perceber que o serviço de um repositório de dados se enquadra num movimento global de ciência aberta e de melhoria do potencial dos dados para a reprodutibilidade da ciência, com a procura de dados mais compatíveis com os princípios FAIR (Fiindable, Accessible, Interoperable, Reusable). Neste aspeto, merece destaque o trabalho colaborativo dos repositórios que usam o software Dataverse, na procura de dotar o software dos requisitos técnicos que facilitem a compatibilidade com os requisitos do CoreTrusSeal e, adicionalmente, documentando o processo e partilhando boas práticas de certificação.

4 CONCLUSÃO

Com esta apresentação esperamos dar uma visão holística do processo de certificação CoreTrusSeal associado a um repositório institucional de uma Universidade, de cariz genérico e que assegura processos amplos que servem uma comunidade diversa do ponto de vista disciplinar. O processo de certificação ganha um maior valor se assegurado como mecanismo de qualificação dos procedimentos e dos fluxos de trabalho, como instrumento de auto-avaliação e melhoria contínua e como ferramenta pedagógica de capacitação profissional. Esta comunicação irá integrar ainda dados dos processos de aprendizagem dos membros da equipa



envolvidos na certificação e os resultados da solicitação do CoreTrustSeal e respetiva validação.

REFERÊNCIAS

CoreTrustSeal (2023). <https://www.coretrustseal.org/about/>

CoreTrustSeal Standards and Certification Board (2022). *CoreTrustSeal Requirements 2023-2025 (V01.00)*. Zenodo.

<https://doi.org/10.5281/zenodo.7051012>

OpenAIRE (2020). *How to find a trustworthy repository for your data*.

<https://www.openaire.eu/find-trustworthy-data-repository>

USDB (2023). *DataRepositoriUM: Para partilhar, publicar e gerir dados de investigação*. *DataRepositoriUM*. <https://datarepositorium.uminho.pt/about/>

USDB (2022). *Guias (in)formativos: Guia do DataRepositóriUM*.

<https://guias.sdum.uminho.pt/datarepositorium>

USDB (2023). *Missão e objetivos do DataRepositóriUM*.

<https://datarepositorium.uminho.pt/about/missao.html>