

## **RODA 3**

### **PRESERVAÇÃO DIGITAL DE LONGA- DURAÇÃO**

CARACTERÍSTICAS E REQUISITOS  
TÉCNICOS

### **SOBRE O DOCUMENTO**

Identificador	WP181215		
Aprovado por	Luís Miguel Ferros	<b>Aprovado em</b>	2018-11-29
Classificação	Público		
Distribuição	N/A		

### **REVISÕES**

<b>#</b>	<b>Data</b>	<b>Autor</b>	<b>Alterações</b>
1	2018-11-29	Luís Miguel Ferros, Miguel Ferreira	Documento inicial

## **SUMÁRIO EXECUTIVO**

O RODA é um repositório digital capaz de incorporar, preservar e dar acesso a todo o tipo de material digital produzido por grandes organizações públicas ou privadas. O seu rol de funcionalidades cobre a totalidade das unidades funcionais do modelo de referência OAIS (*Open Archival Information System*), permitindo que a informação incorporada permaneça autêntica e acessível ao longo do tempo.

Este documento tem como objetivo descrever as principais características, funcionalidades e vantagens do software RODA, bem como os requisitos técnicos necessários à sua correta instalação.

**PRESERVAR E  
DAR ACESSO  
A TODO O  
MATERIAL  
DIGITAL  
PRODUZIDO  
POR GRANDES  
ORGANIZAÇÕES**

**RODA**

O RODA é um repositório digital capaz de incorporar, preservar e dar acesso a todo o tipo de material digital produzido por grandes organizações públicas ou privadas. O seu rol de funcionalidades cobre a totalidade das unidades funcionais do modelo de referência OAIS (*Open Archival Information System*), permitindo que a informação incorporada permaneça autêntica e acessível ao longo do tempo.

O RODA permite gerir informação segundo uma abordagem baseada na análise de risco. O sistema monitoriza permanentemente a informação que detém e alerta o responsável para potenciais riscos que poderão colocar em causa a sua longevidade ou dificultar o seu acesso.

Sendo baseado em standards (OAIS, EAD, METS e PREMIS), este sistema é o parceiro ideal para criar um repositório certificado segundo a norma ISO 16363 (*Audit and certification of trustworthy digital repositories*).

### Compatível com **normas abertas**

O RODA é compatível com normativos como o EAD, EAD 3 e DC para metadados de descrição, PREMIS para metadados de preservação, METS para metadados estruturais e várias normas ao nível dos metadados técnicos (e.g. NISO Z39.87).

### **Autenticidade**

O RODA é suportado por metainformação de preservação (PREMIS) tendo como objetivo assegurar uma cadeia de confiança e autenticidade dos objetos digitais desde o momento que estes entram no repositório. Os metadados de preservação, em conjunto com o estabelecimento da confiança do ambiente circundante (ISO 16363), asseguram a fiabilidade do serviço e a autenticidade dos objetos digitais arquivados.

### **Escalável**

O RODA é baseado numa arquitetura orientada ao serviço, garantindo-lhe assim máxima capacidade para escalar, distribuindo a carga de processamento por todos os servidores necessários para garantir máxima performance. Além disso, o uso de componentes avançados de indexação permite que os serviços de descoberta do RODA sejam distribuídos através de vários servidores num cluster para um desempenho ainda maior.

### **Tecnologia evolutiva**

O mecanismo de plugins suportado pelo RODA permite-lhe ser estendido, sem necessidade de recompilar ou reinstalar o sistema. Os plugins são executados por um avançado motor de execução de tarefas que permite paralelizar a sua execução através de um ou mais servidores.

### **Independente do fornecedor**

O RODA é 100% suportado por tecnologias open-source, sem licenças associadas, permitindo a qualquer instituição assumir responsabilidade pela manutenção do sistema sem depender de um fornecedor específico. Do ponto de vista dos utilizadores, estes apenas necessitam de um browser com suporte para Javascript para que possam tirar partido de todas as funcionalidades do sistema. Isso significa que se podem usar as distribuições de hardware e Linux que melhor se adequam às suas necessidades institucionais.

## ARQUITETURA

O software RODA é composto por um conjunto de módulos aplicativos, os quais estão representados na figura seguinte.

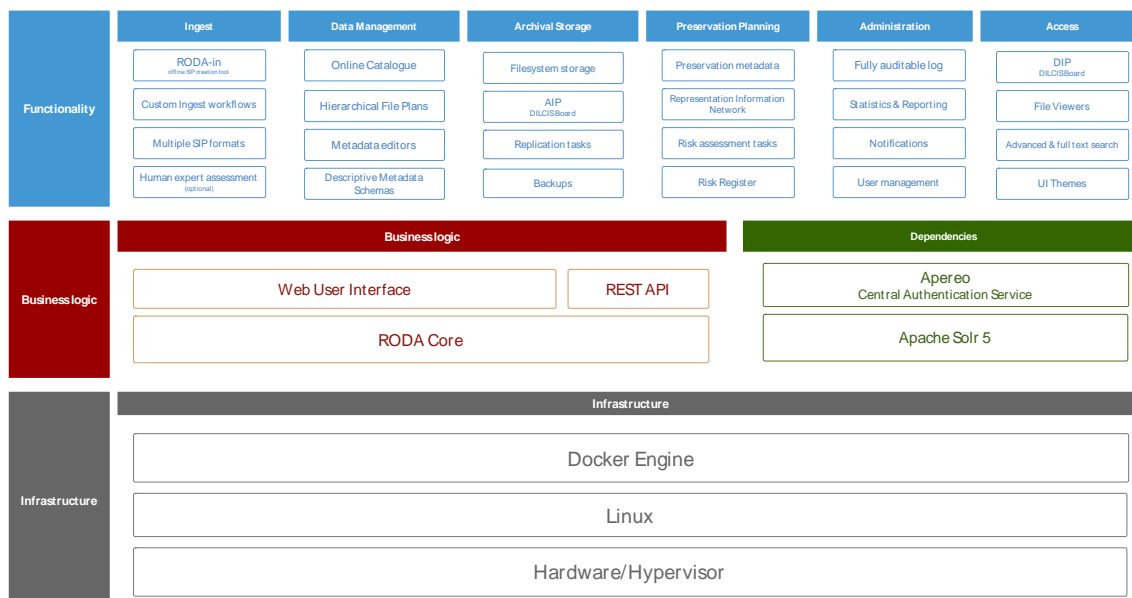


Figura 1 - Arquitetura do RODA.

## ESTRATÉGIAS DE PRESERVAÇÃO APOIADAS PELO RODA

O RODA foi desenvolvido para ser flexível o suficiente para lidar com todas as estratégias de preservação encontradas na literatura. Pode suportar de forma nativa a migração e normalização de formatos, encapsulamento e fornece suporte para emulação.

### MIGRAÇÃO DE FORMATOS

O RODA suporta nativamente a conversão de formatos através do seu sistema de execução de tarefas. A natureza extensível do RODA permite que este seja atualizado para cumprir com novos requisitos e suportar mais e melhores ferramentas de preservação, incluindo conversores.

### ENCAPSULAMENTO

O RODA tem suporte nativo para gestão de Informação de Representação. Este é um conceito-chave que permite documentar tecnicamente a informação de arquivo detida no repositório, permitindo aos utilizadores interpretar a mesma sem ter de recorrer a especialistas ou a fontes de informação externas ao repositório.

## EMULAÇÃO

O RODA é capaz de preservar as versões originais das representações digitais, uma vez que estas foram recebidas através do processo de ingestão. Manter os originais dentro do AIP (Archival Information Package) significa que as estratégias de emulação permanecem viáveis no futuro.

### Ações de preservação incorporadas

As ações de preservação e gestão dentro do RODA são tratadas por um módulo de execução de tarefas. As ações de preservação incluem conversões de formatos, verificações de integridade, diagnóstico de riscos, ações de mitigação de riscos, produção de relatórios, verificações de vírus, etc.

### Suporte para múltiplos formatos

O RODA é capaz de ingerir e normalizar objetos para formatos mais aptos para preservação a longo-prazo: documentos de texto, imagens, bases de dados relacionais, vídeo e áudio. O seu mecanismo de extensão através de plugins permite que o RODA suporte facilmente migrações de formatos adicionais.

### Fluxo de ingestão

O RODA suporta a ingestão de novos materiais digitais, bem como metadados associados em vários formatos distintos. São fornecidas ferramentas para permitir que os produtores criem SIPs (*Submission Information Package*) nos formatos suportados. O fluxo de ingestão pode ser personalizado na interface do utilizador para lidar com políticas e dados especiais.

### Integração com outros sistemas

O RODA expõe as suas funcionalidades através de uma API REST totalmente documentada. Bibliotecas Java estão também disponíveis no GitHub para facilitar a integração com sistemas terceiros.

## Auditável

Todos os utilizadores do RODA têm obrigatoriamente de se registar antes de poderem utilizar o sistema. Todas as ações realizadas por utilizadores ou por processos automáticos ficam registadas para responsabilização futura. As permissões são granulares e podem ser definidas ao nível do repositório ou até mesmo ao nível dos objetos individuais.

A autenticação é suportada por um Serviço de Autenticação Central (CAS), que se pode conectar a vários sistemas de autenticação, como o LDAP, Active Directory, OAuth 1.0/2.0, OpenID, Cartão de cidadão, etc.

## **PLANEAMENTO DA PRESERVAÇÃO BASEADO NO RISCO**

O objetivo de um repositório digital é fazer todos os possíveis para mitigar os riscos que possam impedir a sua capacidade de dar acesso a informação digital de forma autêntica.

A função do repositório é, assim, identificar e avaliar as incertezas que possam afetar o acesso à informação a longo prazo, transformá-las em riscos mensuráveis e implementar ações, através das quais, estes possam ser mitigados efetivamente.

Os riscos que afetam a longevidade da informação não são apenas de natureza tecnológica. Estes podem ser organizacionais, financeiros, relacionados com recursos humanos, etc. Estes podem ainda ser de natureza interna ou externa. Como qualquer organização, os repositórios podem beneficiar de técnicas de gestão de risco para suportar os seus processos de tomada de decisão. A equipa de desenvolvimento do RODA, concordando com este facto, decidiu incorporar a gestão de risco nos processos de gestão do próprio repositório.

O RODA é bastante flexível no que toca aos tipos de dados que podem ser incorporados. O fluxo de ingestão pode ser personalizado para suportar qualquer política institucional, desde as mais restritas, onde um número limitado de formatos e metadados são aceites pelo repositório, até às mais flexíveis, onde o repositório aceita tudo que lhe chegue através dos canais de ingestão. Tirando partido da abordagem baseada na gestão de risco, o gestor do repositório pode decidir, após a ingestão dos dados, qual a melhor ação a ser executada, e não criar constrangimentos artificiais à sua incorporação logo no momento da receção do material.

### **Tarefas de avaliação de risco**

Tarefas de avaliação de risco foram desenvolvidas e são distribuídas sob a forma de plugins. Tipicamente assumem responsabilidades como a deteção de ficheiros para os quais o repositório não apresenta informação de representação suficiente, ficheiros que sofreram adulteração, ficheiros cujo formato está prestes a ser substituído por um novo standard, etc.

Depois de uma avaliação de riscos adequada, o gestor de preservação pode decidir se pretende aceitar o risco ou mitigá-lo através da execução de uma ação de preservação.



## REQUISITOS TÉCNICOS

O produto RODA pressupõe a existência de um conjunto mínimo de dois computadores: um servidor e um posto de trabalho. O computador servidor é responsável por alojar a informação. O posto de trabalho é utilizado pelos utilizadores para administrarem o sistema.

As secções que se seguem descrevem os requisitos mínimos necessários para o correto funcionamento da aplicação nos diversos computadores.

### SERVIDOR

RAM	4 GB 8 GB recomendados
CPU	1.0 Ghz Quad-Core ou superior
Disco	20 GB Depende do número de registos e objetos digitais e da sua taxa de crescimento
Sistema operativo	Ubuntu Server 16.04 LTS ou compatível Sem custos de licenciamento
Software	Docker engine
Rede	100 Mbit/s ou superior 1 Gbit/s é recomendável em contexto de incorporação de grandes volumes de objetos digitais

### POSTO DE TRABALHO

RAM	4 GB
CPU	Intel Dual-Core ou superior
Monitor	1280x768 pixéis ou superior
Sistema operativo	Windows/Linux/macOS
Software	Web browser
Rede	100 Mbit/s ou superior 1 Gbit/s é recomendável em contexto de incorporação de grandes volumes de representações digitais



[www.keep.pt](http://www.keep.pt)



+351 253 066 735



[info@keep.pt](mailto:info@keep.pt)



[sales@keep.pt](mailto:sales@keep.pt)



KEEP SOLUTIONS, LDA.  
Rua Rosalvo de Almeida, n° 5,  
4710-429 Braga  
Portugal

## KEEP SOLUTIONS

A KEEP SOLUTIONS é uma empresa que tem como missão o fornecimento de soluções avançadas para gestão e preservação de informação.

A nossa abordagem assenta no fornecimento de sistemas de informação e na prestação de serviços com vista a permitir aos nossos clientes realizar uma gestão mais eficiente dos seus ativos de informação.

A empresa iniciou a sua atividade em 2008, tendo adquirido o estatuto de spin-off académica da Universidade do Minho, por se tratar de uma iniciativa empresarial com fortes laços de cooperação com centros de investigação e departamentos desta instituição.

Os nossos clientes encontram-se principalmente no setor público, nos domínios cultural, educacional, patrimonial e científico, mais concretamente nas áreas de arquivo, biblioteca e museu.

Apostamos no desenvolvimento contínuo de soluções inovadoras. Para isso, permanecemos ativos na produção de conhecimento científico, participando ativamente em projetos de I&D em cooperação com instituições nacionais e internacionais.