Disponível em www.bad.pt/publicacoes

PAPER

12º CONGRESSO NACIONAL BAD

O repositório digital seguro: diagnóstico e especificação

Ângela Santos[a], *Maria Manuela Pinto*[b]

[a]*Faculdade de Engenharia da Universidade do Porto, Portugal, up201307017@gmail.com*
[b]*Faculdade de Letras da Universidade do Porto, Portugal, mmpinto@letras.up.pt*

Resumo

Esta comunicação resulta de um projeto de dissertação no campo da Ciência da Informação, área da Gestão da Informação (GI), tendo como tema central a criação de um repositório seguro no âmbito da atuação de um provedor de serviços de certificação digital: a MULTICERT, S.A.
O principal objetivo fixado visou contribuir para a estruturação do repositório digital da Multicert desenvolvendo um detalhado diagnóstico organizacional e informacional e consequente elaboração e aplicação de um *Documento de Especificação de Requisitos*.
Sendo a Multicert um provedor de serviços de certificação eletrónica, tem uma grande preocupação com a segurança da informação e pretende que esta se mantenha autêntica, integra, fidedigna, acessível e inteligível no longo prazo. A estruturação de um repositório confiável/seguro constitui, pois, um passo decisivo para a concretização deste objetivo, potenciando a informação como memória, ativo operacional e estratégico, evidência da qualidade do serviço e oportunidade de inovação organizacional. Do estudo realizado, apresentam-se aqui os modelos que sustentaram o diagnóstico da função de GI na empresa, a especificação de requisitos e a autoavaliação do repositório, bem como alguns dos resultados obtidos.

Palavras-chave: Ciência da Informação; Gestão da Informação; Repositório Digital Seguro; Preservação da Informação; Certificação do Repositório

Introdução

A sociedade contemporânea caracteriza-se pelo grande impacto que as Tecnologias de Informação e Comunicação (TIC) têm na vida das pessoas e muito particularmente nas Instituições e demais Organizações.

Algumas das consequências são o crescimento exponencial da produção e fluxo informacional em meio digital, a possibilidade de se armazenarem eletronicamente grandes volumes de informação, quer tivesse sido criada nesse meio ou digitalizada, e a necessidade de a disseminar garantindo a sua preservação e acesso continuado no longo prazo.

É neste contexto que emergem as preocupações com as infraestruturas e ferramentas de armazenamento, preservação e recuperação da informação que se difundem sob a designação mais comum de Repositórios Digitais.

Um Repositório Digital é um termo com uma aceção específica no meio digital, devendo ser sustentável e fiável, como também ser bem enquadrado e bem gerido. Constitui-se como uma plataforma em meio digital na qual é possível aceder à informação em tempo real e sempre que se necessite.

Os repositórios são apresentados de diferentes formas e perspetivas, com uma grande variedade de contextos, comunidades, objetivos e práticas ligadas à sua criação e funcionamento. Podem ser colaborativos e apresentar um baixo controlo dos conteúdos e de acesso aos documentos, como podem ter um alto nível de controlo (Masson, 2008). Contudo, não basta criar um Repositório Digital e aí

colocar toda a informação produzida e recebida pela Organização no decorrer das suas atividades. É necessário garantir a preservação e segurança dessa informação para que possa suportar as Organizações na concretização dos seus objetivos, missão e estratégia. Um repositório deve ter como características fundamentais a reutilização, a acessibilidade, a durabilidade, a flexibilidade, a versatilidade, a interoperabilidade e a inteligibilidade (Sousa, 2013).

Acresce que, nas Organizações a informação é, atualmente, encarada como um dos ativos de maior valia, isto é, uma fonte de vantagem estratégica, exigindo a sua proteção que sejam providenciados os meios necessários para minimizar o risco de esta poder ser "violada" nos seus atributos fundamentais, isto é, em termos de responsabilidade, confidencialidade, autenticidade, integridade, não-repúdio, acessibilidade e inteligibilidade.

É com base na informação, interna e externa, que se tomam as decisões mais importantes e que levam a que a organização se torne competitiva e estratégica e se distinga no meio em desenvolve a sua atividade. Além disso, a complexidade do meio digital e de ciclos de obsolescência de *hardware* e *software* cada vez mais curtos (3 a 5 anos) faz com que a preservação da informação seja indissociável da segurança da mesma e que, ao perspetivar a criação de um repositório seguro/confiável, tenham, também, que ser pensadas, desenvolvidas e implementadas políticas, metodologias, ferramentas e técnicas que garantam a preservação e acesso continuado à informação no longo prazo, no quadro de uma *Política de Preservação e Segurança* alinhada com a *Política de Gestão da Informação* na Organização (ISO16363:2012; Sousa, 2013, 2014; Oliveira, 2014).

No entanto, a operacionalização que está em foco nos diversos instrumentos normativos carece de uma base teórica e metodológica (Silva e Ribeiro, 2002) que oriente a abordagem de cada caso e, na perspetiva da Ciência da Informação, Pinto e Silva (2005) afirmam que as Organizações precisam de "uma abordagem que congregue, desde a fase de conceção da plataforma tecnológica (*hardware* e *software*), até à produção, circulação, avaliação, armazenamento, disponibilização e preservação da informação, toda a Organização e os seus processos de negócio". É com este pressuposto que se procurou contribuir para o desenvolvimento de um SIAP (Sistema de Informação Ativa e Permanente) na Multicert adotando o modelo teórico de base sistémica e informacional com o mesmo nome (Pinto e Silva, 2005) e que visa orientar a análise e a operacionalização da GI no contexto organizacional, envolvendo a Organização e os seus colaboradores no processo de GI com vista a torná-lo mais eficiente e eficaz, o que, no presente estudo, se focou na criação de um repositório digital na em presa Multicert.

Para que o SIAP seja desenvolvido e preservado, por forma a garantir o acesso continuado no longo prazo e a segurança da informação que o integra, é necessário conhecer, compreender e representar o contexto organizacional para depois conhecer, compreender, organizar, representar e gerir o próprio Sistema de Informação (Pinto, 2009).

A abordagem realizada na Multicert que se descreve de seguida poderá ser aplicada a qualquer tipo de Organização. Dar-se-á conta do dispositivo metodológico usado à luz da teoria sistémica e que passou pela adoção do método quadripolar no contexto do modelo SIAP, seguindo-se a apresentação, ao nível do desenvolvimento do pólo técnico, das ferramentas usadas no diagnóstico/autoavaliação e consequente análise da Gestão da Informação (GI) e do sistema tecnológico de *Enterprise Content Management* (ECM) que a sustenta. Incidir-se-á, também, nos referentes orientadores para a certificação do repositório.

1. Metodologia usada

O Método Quadripolar enquadrou e guiou a investigação e o trabalho desenvolvido neste estudo sendo constituído por quatro pólos interatuantes com vista ao desenvolvimento de uma visão holística e dinâmica da abordagem do problema/necessidade em foco e de projetos de operacionalização em permanente avaliação e aperfeiçoamento. Este método é constituído por quatro polos a saber: Pólo Epistemológico, Pólo Teórico, Pólo Técnico e Pólo Morfológico.

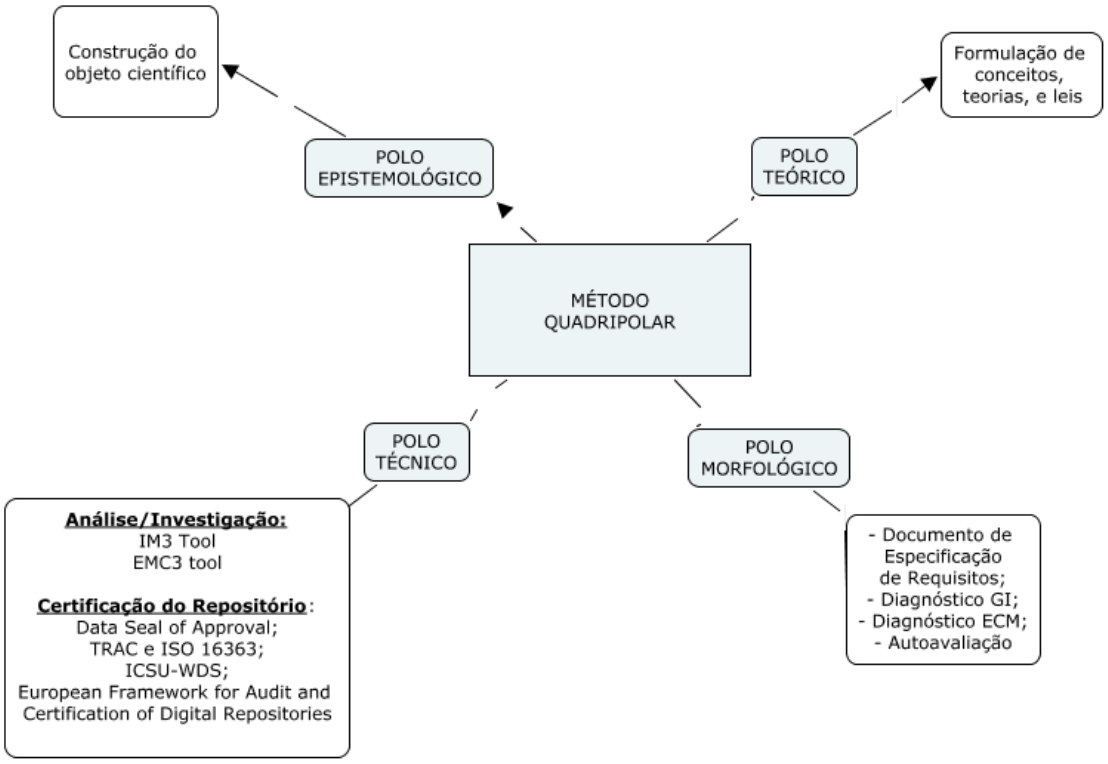

ILUSTRAÇÃO 1 - MÉTODO QUADRIPOLAR APLICADO AO ESTUDO DA MULTICERT

O **Pólo Epistemológico** diz respeito ao debate e à construção do objeto científico, tendo-se definindo aqui os limites da problemática em foco em termos da identificação e delimitação do problema ou necessidade sob o paradigma pós-custodial e científico-informacional que orientou todo o processo.

No **Polo Teórico** manifesta-se a racionalidade que o sujeito reconhece no objeto de estudo, na formulação de conceitos, teorias, e leis. Tomando como referência, o referido paradigma e os conceitos que o mesmo implica, partiu-se da formulação da seguinte questão que serviu para orientar todo o estudo desenvolvido:

- Como pode um fornecedor de produtos e serviços de certificação eletrónica contribuir para a segurança e a preservação da informação numa perspetiva de acesso continuado no longo prazo?

O **Pólo Técnico** compreende a fase onde acontece o contacto com a realidade através de experimentação, observação ou análise/avaliação, com a finalidade de resolver o problema. Sendo este estudo de carácter eminentemente aplicado, o principal enfoque foi qualitativo, acrescendo contributos da investigação-ação dada a forma interativa como se desenvolveu e permitiu a produção e validação de saberes ao longo de todo o processo desenvolvido e no âmbito dos grupos de trabalho que nele participaram. Esta opção visa contribuir tanto para os interesses das pessoas numa situação

problemática imediata, como para promover os objetivos das ciências sociais (O'Brien, 1998). Enquanto o investigador resolve o problema que tem em mãos está, também, a contribuir para o conhecimento científico, divulgando os resultados do seu trabalho. Como refere O'Brien (1998), realizar este duplo objetivo requer a colaboração ativa do investigador com o promotor/cliente, salientando-se, assim, a importância da co-aprendizagem como aspeto fundamental do processo de investigação. Reportamo-nos a um processo simples que se desenvolve de forma cíclica sendo composto por quatro fases - planear, agir, observar e refletir (Kemmis, S. *apud* O'Brien, 1998) – que neste estudo correspondem a ciclos de ajustamentos do mesmo à Organização e problema em foco.

Na fase de planeamento, e como em qualquer estudo, é necessário desenvolver uma revisão da literatura, que partiu de uma pesquisa bibliográfica sobre vários recursos e fontes de informação, nomeadamente artigos, dissertações e instrumentos orientadores e normativos sobre o tema em estudo, e da análise das ferramentas disponíveis para a certificação de repositórios.

Na parte da ação procedeu-se ao levantamento e análise das necessidades da empresa, realizaram-se entrevistas aos Colaboradores da Multicert, observou-se e participou-se ativamente nos diversos grupos de trabalho, recolhendo indicações que conduziram à revisão do estudo e a uma melhor objetivação do problema, passando-se, depois, para ciclos de validação dos instrumentos que se iam desenvolvendo. Foi nesta parte que se recorreu às ferramentas IM3 *tool* e ECM3 *tool* para fazer um diagnóstico sobre o estádio da GI nesta Organização a estes dois níveis.

Para o processo de certificação do repositório, fez-se a análise a ferramentas e projetos que podem servir de ponto de partida a processos similares. Incide-se, de forma particular nos aspetos do diagnóstico/autoavaliação e da certificação do repositório.

No processo de análise e especificação dos requisitos recorreu-se à Engenharia de Requisitos, assumida como a disciplina que se preocupa com a análise (das necessidades) e documentação dos requisitos, fornecendo mecanismos apropriados para facilitar as atividades de análise, documentação e verificação (Lopes, 2004). Em termos operacionais trata-se do processo sistemático de desenvolvimento de requisitos de sistemas, sendo composto por diversas atividades que, na sua execução, recorrem à utilização de um conjunto de técnicas e modelos que tornam sistemática e repetitiva a execução dessas mesmas tarefas (Pinheiro, 2003).

Este processo envolve a identificação dos objetivos do sistema e a descrição das propriedades associadas a restrições ou condicionantes no seu desenvolvimento. O resultado consiste num documento de requisitos onde é descrito o comportamento esperado do sistema em questão (Pinheiro, 2003).

As duas atividades principais são a análise do problema e a especificação de requisitos. Estas duas atividades são interdependentes e devem ser realizadas conjuntamente. A análise e especificação de requisitos envolvem as atividades de determinar os objetos de um sistema e as restrições que lhes estão associadas, assim como, estabelecem o relacionamento entre esses objetivos e restrições e a especificação precisa do sistema, nomeadamente, descrições acerca das propriedades do sistema e descrições associadas a restrições ou condicionantes no seu desenvolvimento. (Pinheiro, 2003).

Este estudo congregou quer a definição da configuração, quer a compreensão e descrição de necessidades e a construção de uma plataforma de informação e comunicação para suportar a estruturação e certificação de um repositório.

As principais operações desenvolvidas neste pólo foram:

- Recolha de bibliografia em fontes diversificadas (bases dados, repositórios);

- Revisão de literatura com base nos métodos de recolha de dados (pesquisa documental, validação de fontes, análise de conteúdo, seleção e síntese com a elaboração de fichas de leitura);
- Identificação e análise de casos similares ao estudo em foco;
- Análise do contexto organizacional da Multicert;
- Observação direta e participante do funcionamento da Multicert e integração em grupos de trabalho;
- Realização de entrevistas exploratórias (aos Colaboradores);
- Levantamento, análise e especificação de requisitos.

No **pólo morfológico** refletiu-se a eficácia destas operações. Aqui assume-se por inteiro a análise dos dados recolhidos e parte-se para a abordagem da problemática em estudo e para a exposição de todo o processo que permitiu a sua construção, nomeadamente na função de comunicação. Trata-se da organização e da apresentação dos dados, devidamente criticados no polo teórico e harmonizado no polo epistemológico, o que ilustra o peso interativo da investigação quadripolar.

Neste pólo foram analisados os dados da avaliação da "maturidade" da GI e do ECM, o Sistema de Gestão de Documentos, neste caso o ALFRESCO.

Desta forma, sistematizaram-se e apresentaram-se os resultados obtidos através de um Documento de Especificação de Requisitos com vista à conformidade e eventual certificação do Repositório. O documento de requisitos aparece estruturado em três grandes secções: Infraestrutura Organizacional; Gestão de Objetos Digitais; Infraestrutura e Gestão de Riscos de Segurança. Este documento está, também, dividido em requisitos funcionais e não funcionais. Os requisitos funcionais descrevem os serviços que o sistema deve fornecer, especificando como este deve reagir por exemplo a solicitações dos utilizadores (Pinheiro, 2003). Dizem respeito à definição das funções que um sistema ou um componente de um sistema deve fazer. Descrevem as funções a realizar nas entradas de um sistema ou em um dos seus componentes, a fim de que se produzam saídas (Lopes, 2004).

Neste pólo apresentam-se, ainda, os resultados relativos às análises referidas no pólo técnico e que são apresentadas através do *Diagnóstico GI*, do *Diagnóstico ECM* e da *Autoavaliação*, segundo o instrumento normativo ISO 16363.

2. INSTRUMENTOS USADOS NO DIAGNÓSTICO

A GI apresenta-se como uma área de estudos aplicada e uma função fundamental para qualquer Organização no desenvolvimento da sua atividade.

Neste estudo assume-se a GI como o «[…] estudo, conceção, implementação e desenvolvimento dos processos de gestão inerentes ao fenómeno infocomunicacional, incluindo a identificação, compreensão, representação lógica e redesenho dos processos organizacionais e configurações físicas e/ou meios tecnológicos que modelam a produção, fluxo, uso, disseminação e preservação da informação, no contexto da ação humana e social» (Pinto, 2014a).

Inserem-se aqui os processos através dos quais a Organização planeia, cria, recebe, adquire, organiza, armazena, protege, divulga e preserva a sua informação. É, também, o meio através do qual a Organização garante que o valor dessa informação é identificado e explorado. O principal objetivo é

garantir a eficiência e eficácia dos referidos processos para que a informação correta esteja disponível para a pessoa certa, no momento certo, no formato e meio certo, independentemente do espaço e tempo em que esta se encontre.

A "maturidade" da atividade de GI é considerada um fator-chave para o sucesso, constituindo o suporte dado pela plataforma tecnológica uma peça fundamental do todo organizacional., existindo diversos instrumentos de apoio no sentido de uma maior qualidade do processo, serviço e resultado final, destacando-se entre eles, os designados "Modelos de Maturidade" (Public Record Office Vitoria, 2014).

Estes modelos de diagnóstico fornecem um poderoso instrumento para a determinação do estádio em que se encontra a organização (estratégias e práticas) e para o planeamento das ações necessárias para progredirem e, com isso, alcançarem os objetivos desejados. Baseiam-se na premissa de que pessoas, organizações, áreas funcionais, processos, etc., evoluem através de um processo de crescimento que passa por um determinado número de estádios distintos (Public Record Office Vitoria, 2014).

Para a avaliação da maturidade da GI na Multicert selecionou-se o IM3 *tool* [1] (*Information Management Maturity Measurement Tool*), uma ferramenta de auto-avaliação aplicada aos organismos governamentais australianos e cujos objetivos passam por:

- Identificar pontos fortes e fracos na GI;
- Priorizar áreas de GI que precisam de mais atenção;
- Auxiliar no estabelecimento de metas para a capacitação e as habilidades de desenvolvimento em GI;
- Apoiar iniciativas para melhorar a GI.

Não se trata de uma lista de verificação de conformidade, mas de uma ferramenta de autoavaliação para determinar a estratégia, planeamento e práticas de GI na Multicert, com o foco na organização, nos seus processos, atores e informação. O seu caráter abrangente não suscitou problemas à aplicação em contexto empresarial.

Esta ferramenta encontra-se dividida em 4 secções que respondem a diferentes questões acerca de:

- Pessoas (conhecimento, habilidade, experiência e atitude dos colaboradores de maneira a contribuir para uma boa GI);
- Organização (papel da GI na Organização);
- Ciclo de vida e qualidade da Informação (gestão de ativos de informação específicos na Organização, com vista a um acesso a longo prazo a informação de qualidade);
- Sistemas e processos de negócio (sistemas e processos eletrónicos e manuais que apoiam as práticas de GI na Organização).

Contém 17 questões cada uma das quais tem associadas 7 respostas, estando as primeiras 5 numeradas de 1 a 5, correspondendo o 1 ao menos satisfatório e o 5 ao mais satisfatório:

[1] PUBLIC RECORD OFFICE VITORIA. Information Management Maturity Measurement Tool – IM3 [Em linha]. North Melbourn. [Consult. 21 Nov. 2014] Disponivel na Internet:<URL:HTTP:// http://prov.vic.gov.au/government/information-management/information-management-maturity-measurement-tool-im3

1. *Não gerido*
2. *Consciente*
3. *Em desenvolvimento*
4. *Operacional*
5. *Proactivo*

Necessita de mais informação

Não aplicável.

O objetivo passa pela atribuição de um "nível" (de acordo com a escala pré-definida) que melhor se adeqúe à realidade da organização para assim se ter uma noção real de como a GI está a ser desenvolvida em toda a Organização e como esta:

- Cria, adquire e troca informação;
- Gere, adquire e atribui a propriedade da informação;
- Fornece acesso à informação interna e externa;
- Assegura e verifica a qualidade da Informação;
- Identifica o valor da informação e utiliza esse valor;
- Dá suporte aos colaboradores para concretizar as suas responsabilidades de GI.

No que concerne à maturidade do sistema tecnológico ECM, utilizou-se o ECM *maturity model*[2]. Este centra-se no sistema tecnológico de suporte e tenta fornecer um quadro estruturado para a construção de um roteiro no contexto de uma estratégia global.

Esta *framework* sugere níveis de capacidade graduados, que vão desde a recolha de informação rudimentar e controlo básico através de níveis cada vez mais sofisticados de gestão e integração, resultando num estado de maturação de contínua experimentação e melhoria.

Como todos os modelos de maturidade, este é igualmente descritivo e prescritivo, podendo ser aplicado para auditar, aceder e explicar o estado corrente do ECM, bem como fornecer um *roadmap* para desenvolver as capacidades, neste caso, da Multicert.

O modelo comporta 13 dimensões distribuídas em 3 categorias: *Pessoas*, *Informação* e *Sistemas*. Estas dimensões devem aplicar-se a qualquer empresa, independentemente do setor, tamanho, tecnologia e objetivos de negócio.

A categoria *Pessoas* relaciona-se com as pessoas e todos os atributos que melhoram a maturidade, tais como o conhecimento de diferentes tipos de profissionais dentro de uma Organização, bem como as interações para assegurar o alinhamento estratégico com os objetivos institucionais para o sucesso. Esta categoria possui quatro dimensões de maturidade: especialistas em TI, especialistas no negócio, processo e alinhamento estratégico. A dimensão de maturidade *especialistas no negócio* diz respeito ao empregado, à educação executiva e à compreensão dos preceitos de ECM. A de *especialistas em TI* implica a capacidade de aproveitar adequadamente os novos sistemas. A de *Processo* é a profundidade com que a empresa analisou os seus processos de negócio, com uma orientação para o conteúdo informacional. A de *Alinhamento estratégico* refere-se à extensão.

[2] PELZ-SHARPE, Alan, et al. (2009) – ECM3: Ecm Maturity Model. [Em linha]. V. 1.0. [Consult. 03 Mar. 2015]. Disponível na Internet: <URL: https://ecmmaturity.files.wordpress.com/2009/02/ec3m-v01_0.pdf>

A categoria *Informação* refere-se aos atributos que afetam o conteúdo no aplicativo, incluindo a capacidade de gerir o conteúdo. Esta categoria tem cinco dimensões: conteúdo/meta-informação, profundidade, políticas, reutilização e capacidade de pesquisa. A dimensão de *conteúdo ou meta-informação* consiste na extensão da análise de conteúdo e meta-informação. A de *profundidade* diz respeito à completude da gestão do ciclo de vida do conteúdo. A dimensão *Políticas* refere-se à extensão das políticas e procedimentos de GI. A *reutilização* incide no conteúdo existente em aplicativos ECM que pode ser reutilizado para outros fins que não o que lhe estavam inicialmente destinado. A dimensão *capacidade de pesquisa* refere-se à capacidade de encontrar o conteúdo certo no momento certo.

A Categoria *Sistemas* diz respeito à capacidade técnica do aplicativo do negócio, a eficaz / colaboração de TI, a compreensão e a sincronização. Para esta categoria são especificadas quatro dimensões: alcance, abrangência, segurança e usabilidade. A dimensão *Alcance* diz respeito às características funcionais relevantes de ECM que foram adotadas (por exemplo, gestão de documentos, gestão de processos de negócios, gestão de ativos digitais, etc.). A da *abrangência* refere-se à evolução do departamento de sistemas de gestão empresarial sempre que necessário. A da *segurança* é a medida em que o acesso a conteúdo real reflete os direitos da empresa e a *usabilidade* comporta a aptidão para os fins do aplicativo ECM.

Tal como na avaliação com o IM3 *tool*, o objetivo passa por fazer corresponder a realidade da Organização com o quadro pré-definido, e que também está dividido por categorias de maturidade, e, assim, perceber-se como está o ECM a responder aos diferentes objetivos.

Trata-se de duas análises complementares testadas no desenvolvimento deste estudo e que requerem, num futuro próximo, uma análise ao nível de modelos que ultrapassa os objetivos do estudo realizado.

No entanto, e para já, demonstrou a possibilidade de serem equacionados como referentes de diagnóstico a utilizar ao nível do pólo técnico do dispositivo metodológico quadripolar, tal como acontece com outros referentes normativos ou de boas práticas como uma ISO 15489, ISO 30300, ISO 30301, entre muitas outras. Não sendo possível, por questões de confidencialidade, discutir resultados concretos, ficou patente uma visão mais alargada sobre a GI na Multicert e o contributo que as plataformas tecnológicas estão a dar para uma Gestão mais eficaz, constituindo um importante referente de análise para se avançar para pensar a estruturação de um repositório digital confiável e uma futura Certificação.

3. Repositórios Digitais Confiáveis: conceito e modelo

A propósito da GI, e mais especificamente ao nível da preservação e segurança da informação no contexto de utilização das TIC, surge o repositório como componente imprescindível e contribuinte para a implementação do processo de preservação, assumida esta como variável da GI (Pinto, 2009, 2014b).

Neste estudo a segurança e a preservação da informação incidem no repositório digital, mas, como se pode verificar, este estudo não poderia ser olhado sem abordar a Organização Multicert, as suas políticas, estratégias, processos, atores e, sobretudo, o sistema de informação organizacional (SIO) a proteger e preservar de forma ativa e permanente assumindo-o como sendo «[…] constituído pelos diferentes tipos de informação registada ou não externamente ao sujeito, não importa qual o suporte, de acordo com uma estrutura prolongada pela ação na linha do tempo» (DeltCI, 2014) e distinto do Sistema Tecnológico de Informação (STI) que «[…] suporta o fenómeno e processo

infocomunicacional, permite uma comunicação assíncrona e multidireccionada e potencia o acesso à informação» (Pinto, 2014).

Os repositórios são apresentados de diferentes formas e perspetivas, com uma grande variedade de contextos, comunidades, objetivos e práticas ligadas à sua criação e funcionamento, podem ser colaborativos e apresentar um baixo controlo dos conteúdos e de acesso aos documentos, como podem ter um alto nível de controlo (Masson, 2008).

Para alguns autores constituem uma "ferramenta de TI" que permite o armazenamento da informação produzida no decorrer das funções de qualquer Organização e, de acordo com o Infonet[3], um repositório digital é um "meio" de gestão, armazenamento e fornecimento de acesso a conteúdo digital que deve ser sustentável, fiável, bem enquadrado e bem gerido.

Os repositórios não se destinam apenas à produção científica ou ao armazenamento de informação de carácter pedagógico, podendo qualquer Organização ter necessidade de criar o seu repositório de informação organizacional produzida em meio digital.

Um repositório digital é, assim, uma plataforma (*hardware* e *software*) que serve para armazenar, preservar e difundir informação, suportando diversos processos e serviços. Inclui uma componente tecnológica (ao nível do STI), uma componente informacional (ao nível do SIO) e uma componente de serviço que são indissociáveis. Apresenta como características fundamentais a reutilização, a acessibilidade, a durabilidade, a flexibilidade, a versatilidade, a interoperabilidade e a inteligibilidade, mas também garantir a confidencialidade, a autenticidade, a integridade e o não-repúdio da informação nele contida.

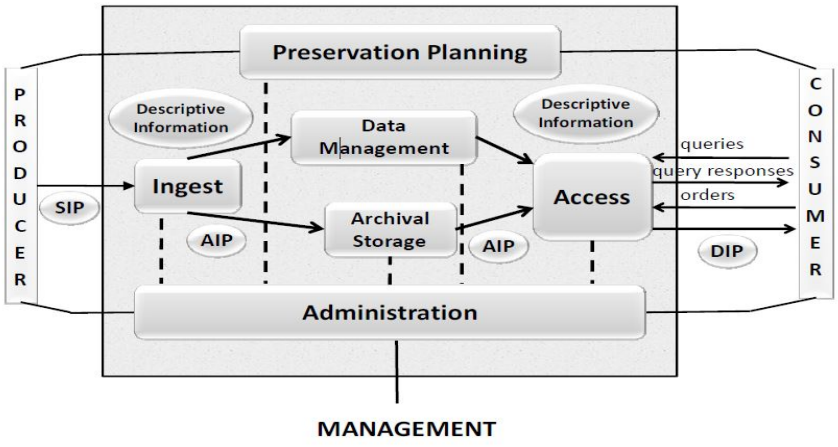

ILUSTRAÇÃO 2 MODELO CONCEPTUAL OAIS (CCSDS, 2012)

Para a sua criação é fundamental o modelo conceptual que o sustenta sendo o Modelo de referência OAIS[4] o mais utilizado a nível mundial. Este parte da definição de regras e linhas orientadoras para sensibilizar e ampliar a consciência e compreensão de conceitos, definir terminologias para descrever e comparar modelos de dados e arquiteturas de arquivo, promovendo o consenso sobre os elementos e os processos relacionados com a preservação e acesso à informação digital (CCSDS, 2012). Trata-se de um esquema básico e uma ferramenta para a melhoria da comunicação e produtividade entre os diversos níveis e as diferentes comunidades, sendo identificado como um Sistema Aberto de Arquivo de Informação (SAAI) (Backer, et. al, 2009) e considerando os arquivos como "organizações de pessoas e sistemas" que aceitaram essa responsabilidade de preservar e comunicar a informação

[3] JISC INFONET. Infokits. *Repositórios* . [Em linha]. [s.l]. [Consult. 21 Abr. 2015]
Disponível na Internet: <URL:http:// www.jiscinfonet.ac.uk/infokits/repositories>

(CCSDS, 2012). O modelo prevê três entidades - o produtor, o consumidor, e o administrador – e contém um modelo de informação para inserção da meta-informação colocando o foco na preservação.

4. Referenciais para a certificação do Repositório

Seja qual for a plataforma de suporte ou as estratégias de preservação adotadas, um repositório deve corresponder às expectativas criadas pelos seus utilizadores. Afirmar que se é capaz de garantir o acesso continuado à informação digital não é suficiente para estabelecer um clima de confiança junto dos vários intervenientes que interagem com o repositório, i.e. produtores, consumidores, operadores do sistema, gestores, entidades de fomento, ou outros, e para que seja verdadeiramente confiável, é fundamental que existam formas de medir e demonstrar essa confiabilidade (RLG et al., 2007).

O repositório deve acima de tudo ser seguro, não se limitando apenas aos aspetos tecnológicos mas também às instalações físicas e às ações pessoais. Estes devem incluir uma análise sistemática da informação, sistemas, pessoas e instalações físicas; adotar procedimentos de controlo para tratar adequadamente as necessidades de segurança; delinear papéis, responsabilidades e autorizações relativas à implementação de mudanças no sistema; planear as ações de preservação e reparação, a prevenção de desastres, entre outros aspetos, abarcando tudo o que é mantido no repositório (informação, meta-informação, registos de ações realizadas, pistas de auditoria etc.).

A preservação é, pois, uma condição *sine qua non* para garantir a qualidade, partilha e uso da informação no longo prazo em repositórios digitais sustentáveis. A informação criada precisa ser gerida, preservada e arquivada de modo a que responda às necessidades e memória organizacional e que os elevados investimentos na sua produção, preparação e apresentação não sejam perdidos.

A certificação é fundamental para garantir a confiabilidade dos repositórios digitais e, portanto, para sustentar as oportunidades para o acesso e uso da informação no longo prazo e serviços correspondentes.

Nos últimos anos, têm sido desenvolvidas em todo o mundo uma série de normas para a certificação e procedimentos de acreditação, nomeadamente: a norma ISO 16363:2012 e o TRAC (Trustworthy Repository Audit Certification); o *Data Seal of Approval* (DSA); a certificação WDS - Trusted Data Services for Global Science - do International Council for Science - World Data System (ICSU-WDS).

Neste contexto, e no que respeita a uma visão global do estudo desenvolvido, e dado constituir uma proposta recente no âmbito internacional, foi considerado, ainda, o *European Framework for Audit and Certification of Digital Repositories*.

A criação do *European Framework for Audit and Certification of Digital Repositories* em 2010, direcionado à auditoria e certificação de Repositórios Digitais, resulta da congregação de esforços de três entidades que trabalham no âmbito da certificação de repositórios e que pretendem que sejam criados mecanismos para garantir que os grupos possam colaborar na criação de um quadro integrado para auditoria e certificação de repositórios digitais.

- o CCSDS (Consultative Committee for Space Data Systems)/ISO Repository Audit and Certification Working Group (RAC);
- o Data Seal of Approval Board;
- e o DIN Working Group "Trusted Archives - Certification".

[4] Foi publicado em 2002 pelo Consultative Committee for Space Data Systems (CCSDS) e adotado como norma em 2003 (ISO 14721:2003).

Esta *framework* é composta por uma sequência de três níveis de certificação, correspondendo ao aumento da confiabilidade obtida:

1. a certificação básica (*basic certification*) que é concedida aos repositórios que obtêm certificação Data Seal of Approval (DSA) [5];
2. a certificação alargada (*extended Certification*) que é concedida aos repositórios com a certificação básica e que realizam adicionalmente uma autoavaliação estruturada, com revisão externa e disponível publicamente, baseadas na ISO 16363, ou na DIN 31644;
3. a certificação formal (*formal certification*) que é concedida aos repositórios que obtêm, além da certificação básica, uma auditoria externa total e certificação conforme à ISO 16363 ou 31644 DIN.

A concessão desses certificados permitirá aos repositórios mostrar um dos três símbolos nas suas páginas web e outros documentos, além de quaisquer outras marcas de certificação da DSA, ISO ou DIN.

O DSA e o WDS oferecem uma certificação básica para repositórios digitais confiáveis. A sua especificação de requisitos e procedimentos de avaliação baseiam-se nos mesmos princípios de abertura e transparência, no sentido de um justo equilíbrio entre a simplicidade e robustez do trabalho e esforço envolvidos.

O DSA, criado em 2009, constitui-se como um selo de aprovação de dados para repositórios digitais que garante que a informação arquivada possa ser encontrada, compreendida e utilizada no futuro.

As Diretrizes de qualidade formuladas para o DSA são de interesse para os produtores e instituições que criam e trabalham com "dados" digitais, para as organizações que armazenam informação, e para os consumidores de informação, nomeadamente dados científicos. Os objetivos do Selo de Aprovação de Dados são: proteger os dados, garantir uma alta qualidade e orientar o manuseamento de dados confiáveis no futuro sem a necessidade da implementação de novas normas, regulamentos ou custos elevados.

O DSA visa conferir aos produtores a garantia de que a sua informação e materiais associados serão armazenados de forma fiável e podem ser reutilizados, proporcionando aos organismos de financiamento confiança de que os dados continuarão disponíveis para reutilização e os seus investimentos se perderão. Dá também suporte aos repositórios de dados. Trata-se, portanto, de um conjunto de boas-práticas que se pretendem que sejam seguidas sobretudo por organizações responsáveis pela preservação de dados científicos.

É composto por 16 requisitos baseados em 5 critérios: que os dados podem ser encontrados na internet; que estão acessíveis; que estão num formato usável; que são confiáveis; e que são identificados num modo único e persistente. Dos 16 requisitos, 3 dizem respeito aos produtores e ao processo de ingestão, 10 à qualidade do repositório e 3 ao acesso à informação por parte dos consumidores. O cumprimento de cada requisito é avaliado numa escala de 0-4.

Por sua vez, o ICSU – WDS[6] foi criado em outubro de 2008. Os objetivos do sistema de dados WDS são: preservar a qualidade de dados científicos e outra informação; facilitar o acesso aberto; e promover a adoção de normas. Este sistema baseia-se no potencial oferecido pelas interligações entre

[5] DATA ARCHIVING AND NETWORKED SERVICES. Data Seal of Approval [Em linha]. [s.l] [Consult. 20 Mar. 2015] Disponivel na Internet: <URL: http://www.datasealofapproval.org/en/>

[6] INTERNATIONAL COUNCIL OF SCIENCE. World Data System (WDS) [Em linha]. [s.l]. [Consult 21 Mar. 2015] Disponivel na Internet:<URL: http://www.icsu.org/what-we-do/interdisciplinary-bodies/wds/?icsudocid=about

os componentes avançados de gestão de dados para promover aplicações disciplinares e multidisciplinares para o benefício da comunidade científica internacional e outras partes interessadas.

Até ao momento, as duas normas têm evoluído e operado de forma independente. Apesar de afirmarem missões totalmente multidisciplinares, o foco principal da DSA dirige-se a repositórios digitais nas Ciências Humanas e Sociais e, por razões históricas, o foco do ICSU-WDS são as Ciências da Terra e do espaço. .

Embora as duas organizações tenham diferenças consideráveis, uma parceria promoveria ganhos de eficiência, simplificação de opções de avaliação, estimulando mais certificações, e aumentando o impacto na comunidade e como um possível primeiro passo para o desenvolvimento de um quadro comum para a certificação e um serviço de repositórios de dados confiáveis.

No que concerne aos modelos de requisitos de auditoria e certificação é de referenciar o TRAC (RLG et al., 2007). Este consiste num conjunto de requisitos que vão desde a gestão organizacional às infraestruturas de suporte que visam assegurar a confiança em torno de um repositório. Trata-se de uma ferramenta que permite auditar, avaliar, e certificar os repositórios digitais, identificando para isso a documentação necessária para realizar uma auditoria e estabelecendo metodologias apropriadas para determinar a robustez e a sustentabilidade do repositório digital. Depois de anos de desenvolvimento em torno da *checklist* do TRAC surge a norma ISO 16363:2012.

A ISO 16363:2012, apresenta como objetivo essencial fornecer uma lista de requisitos a partir da qual se possam realizar auditorias aos repositórios no sentido do controlo de ameaças e riscos à sua segurança com base em 109 critérios, apresentando a sua *checklist* quer requisitos de preservação, quer requisitos de segurança. Esta norma constituiu a principal referência do estudo realizado na Multicert, complementando a abordagem e certificações já efetuadas pela Multicert na perspetiva das tecnologias e com recurso às normas da série ISO 27000 e orientações ao nível da Gestão do Risco.

5. Resultados

Com os resultados dos diagnósticos realizados e da análise das ferramentas para a certificação do repositório, procedeu-se à elaboração do *Documento de Especificação de Requisitos* para o repositório digital da Multicert, em linha com estudos como o de Sousa (2013, 2014) e de Oliveira (2014) e tendo como base a norma ISO 16363:2012.

Com este *Documento* procedeu-se ao diagnóstico de conformidade do ALFRESCO aos requisitos da norma ISO 16363:2012, na perspetiva do repositório (numa das arquiteturas previstas no MoReq2010: "*in place" records management* ou "*in app*" *records management*), seguindo-se a elaboração de documentos fundamentais para atingir a conformidade nos casos de requisitos parcialmente cumpridos e não cumpridos.

A norma ISO 16363 é, de facto, o instrumento de referência quer para uma futura certificação e consequente monitorização, planeamento e manutenção, quer para a implementação das estratégia e ações a levar a cabo no âmbito da missão fixada para o repositório digital confiável da Multicert, no contexto de uma gestão (e preservação) de informação cada vez mais colaborativa.

Para manter o estado de confiança, o repositório deve realizar frequentemente auditorias e comunicá-las ao seu público-alvo que são essencialmente os colaboradores internos da Multicert. Ao realizar frequentemente as auditorias ao repositório e divulgá-los aos principais interessados (Colaboradores), estes vão criar um "laço" de confiança com a empresa, os processos que desenvolve e os

recursos/plataformas que utiliza, incentivando-os a potenciar o recurso "repositório" de uma forma cada vez mais antecipada e abrangente.

Com a elaboração deste estudo tornou-se claro que a Multicert está no bom caminho para a Certificação do Repositório. O primeiro passo foi dado pois fez-se o diagnóstico necessário para perceber como as coisas estão a ser feitas atualmente e aquilo que se pode e deve fazer-se para o caminho da certificação.

Conclusões

Afirmar que se é capaz de garantir o acesso continuado à informação digital não é suficiente para estabelecer um clima de confiança junto dos vários intervenientes que interagem com o repositório, i.e. produtores, consumidores, operadores do sistema, gestores, entidades de fomento, ou outros. Para que um repositório seja verdadeiramente confiável, é fundamental que ao ser concebido se considere a existência de formas de medir e demonstrar essa confiabilidade (RLG et al., 2007).

O levantamento e análise dos instrumentos de diagnóstico e de certificação com vista à criação de um repositório digital de confiança (e seguro) constitui um passo decisivo e integra os resultados obtidos no estudo realizado na MULTICERT.

De facto, o repositório abrange aspetos informacionais, tecnológicos / físicos, bem como decisões e ações organizacionais e individuais. Seja qual for a plataforma de suporte ou as estratégias de preservação adotadas, um repositório deve corresponder às expectativas criadas pelos seus utilizadores e, sobretudo, prossumidores.

Um repositório digital confiável, e consequente certificação, constitui a via que as Organizações dispõem para garantir os principais atributos da informação em meio digital e prover à preservação e acesso continuado à mesma pelo tempo necessário, seja na perspetiva da empresa, seja dos seus clientes e, em última instância, da comunidade em que se insere.

Para um provedor de serviços de certificação eletrónica é vital assegurar a segurança da informação e garantir a sua autenticidade, integridade, fidedignidade, não-repúdio, acessibilidade e inteligibilidade no longo prazo. A estruturação de um repositório confiável/seguro é, pois, uma prioridade que acabará por potenciar a informação como ativo operacional e estratégico, como evidência da qualidade dos serviços que presta, oportunidade de inovação organizacional e memória da própria empresa.

Referências bibliográficas

CONSULTATIVE COMMITTEE FOR SPACE DATA SYSTEMS (CCSDS) (2012) Reference model for an Open Archival Information System (OAIS). Magenta Book. [Consult. 23 Nov. 2014]. Disponível na Internet:<URL: http://public.ccsds.org/publications/archive/650x0m2.pdf

DATA ARCHIVING AND NETWORKED SERVICES . Data Seal of Approval [Em linha]. [s.l] [Consult. 20 Mar. 2015] Disponível na Internet: <URL: http://www.datasealofapproval.org/en/>

INTERNATIONAL COUNCIL OF SCIENCE. World Data System (WDS) [Em linha]. [s.l]. [Consult 21 Mar. 2015] Disponivel na Internet:<URL: http://www.icsu.org/what-we-do/interdisciplinary-bodies/wds/?icsudocid=about

ISO 16363:2012. Space data and information transfer systems : Audit and certification of trustworthy digital repositories. Geneva, Switzerland: ISO, 2012

LOPES, Luís (2004) - Um Modelo de Processo de Engenharia de Requisitos para Ambientes de Desenvolvimento Distribuído de Software. PUCRS: Faculdade de Informática, Porto Alegre. Dissertação de Mestrado

JISC INFONET. Infokits. Repositórios . [Em linha]. [s.l]. [Consult. 21 Abr. 2015]

Disponível na Internet: URL:http:// www.jiscinfonet.ac.uk/infokits/repositories>

MASSON, Sílvia M. (2008) – "Os Repositórios digitais no âmbito da Sociedade Informacional". Prisma.com [Em linha]. Nº 7. [Consult. 03 Mar. 2015]. Disponível na Internet: <URL: http://prisma.cetac.up.pt/105_Repositorios_digitais_no_ambito_da_Sociedade_Informacional_Silvia_Masson.pdf >

O'BRIAN, Rory (1998) - An Overview of the Methodological Approach of Action Research. Faculty of Information Studies, University of Toronto. [Em linha]. [Consult. 03 Mar. 2015]. Disponível na Internet: <URL: http://www.web.net/~robrien/papers/arfinal.html>

OLIVEIRA, Hugo (2014). A Preservação da informação : um contributo para a implementação de um arquivo digital certificável no Município do Porto. Porto: Faculdade de Engenharia da Universidade do Porto, 2014. Orientadora: Maria Fernanda Martins; coorientadora: Maria Manuela Pinto. Dissertação de Mestrado em Ciência da Informação

PELZ-SHARPE, Alan, et al. (2009) – ECM3 : Ecm Maturity Model. [Em linha]. V. 1.0. [Consult. 03 Mar. 2015]. Disponível na Internet: <URL: https://ecmmaturity.files.wordpress.com/2009/02/ec3m-v01_0.pdf>

PINHEIRO, Olga (2003). Sistema de Apoio à Decisão no Planeamento da Produção de Produtos Complexos: Identificação e Especificação de Requisitos. Porto: Faculdade de Engenharia da Universidade do Porto. Dissertação de Mestrado.

PINTO, Maria Manuela; SILVA, Armando M. (2005) - Um modelo sistémico e integral de gestão da informação nas organizações. In CONTECSI - CONGRESSO INTERNACIONAL DE GESTÃO DA TECNOLOGIA E SISTEMAS DE INFORMAÇÃO, 2º, São Paulo, 2005 – Actas do congresso. [CD-ROM]. São Paulo: TECSI-FEA-USP, 2005. [Consult. 3 de Nov. 2014]. Disponível na Internet: <URL: http://ler.letras.up.pt/uploads/ficheiros/3085.pdf>

PINTO, Maria Manuela (2009) - PRESERVMAP : Um roteiro da preservação na Era Digital. Porto: Edições Afrontamento; CETAC.Media (Colecção CAI; 7). ISBN 978-972-36-1070-3

PINTO, Maria Manuela (2014a). "Da Preservação de Documentos à Preservação da Informação". In DUARTE, Zeny - A conservação e a restauração de documentos na era pós-custodial. EDUFBA - Editora da Universidade Federal da Bahia, 2014, p.127-196. ISBN: 978-85-232-1240-7.

PINTO, Maria Manuela (2014b) - Gestão e Preservação da Informação : o impacto do pensamento sistémico. In Encontro Internacional de Arquivos. Évora : Universidade de Évora, 2014.

PUBLIC RECORD OFFICE VITORIA. Information Management Maturity Measurement Tool – IM3 [Em linha]. North Melbourn. [Consult. 21 Nov. 2014] Disponivel na Internet:<URL:HTTP:// http://prov.vic.gov.au/government/information-management/information-management-maturity-measurement-tool-im3.

RESEARCH LIBRARY GROUP; NATIONAL ARCHIVES AND RECORDS ADMINISTRATION; ONLINE COMPUTER LIBRARY CENTER (2007). Trustworthy repositories audit & certification : Criteria & Checklist. [Em linha].[Consult. 23 Nov. 2014], Disponivel na Internet:<URL: http://www.crl.edu/sites/default/files/attachments/pages/trac_0.pdf >

SILVA, Armando M.; RIBEIRO, Fernanda (2002) - Das "ciências" documentais à ciência da informação : ensaio epistemológico para um novo modelo curricular. Porto: Edições Afrontamento, 2002. ISBN: 972-36-0622-4.

SOUSA, Paula (2014).- Gestão da Informação: do modelo de segurança e preservação ao repositório confiável. Páginas a&b: arquivos & bibliotecas. Lisboa. ISSN 0873-5670. S. 3, 1 (2014) 91-119. [Em Linha]. [Consult. 18 dez. 2014]. Disponível em www:<url: http:// http://ojs.letras.up.pt/index.php/paginasaeb/article/view/572>.