

Novos caminhos para as bibliotecas e bibliotecários de saúde: curadoria de dados, cloud e web semântica

New trends for health libraries and librarians: data curation, cloud and semantic web

Paula S. SARAIVA. Universidade de Évora, Portugal. (pcs.saraiva@gmail.com)

Paulo QUARESMA. Universidade de Évora, Portugal. (pq@di.uevora.pt)

Resumo

O 3º milénio é a era da sociedade da informação e das transformações proporcionadas pela aplicação das tecnologias ao quotidiano das pessoas. As bibliotecas, nomeadamente as da área da saúde, onde a investigação científica exige informação cada vez mais precisa, credível e atualizada, necessitam de se adaptar constantemente às mudanças tecnológicas para oferecerem novos serviços que melhor se adaptem ao perfil dos seus utilizadores. Atualmente, as bibliotecas são produtoras e difusoras de conhecimento: produzem conteúdos, colaboram em rede, disponibilizam conteúdos e aplicações para dispositivos móveis, fornecem serviços à distância e preservam os dados digitais para que mais tarde possam ser reutilizados, gerando novo conhecimento. É neste contexto que os bibliotecários têm que desenvolver novas competências para oferecer novos serviços, aprender novas formas de gerir as suas coleções (agora também elas digitais) e recuperar a informação de forma mais rápida e eficiente. Esta comunicação dá ênfase às novas tendências e caminhos que as bibliotecas terão que seguir, num futuro (que já é presente) e onde conceitos como a web semântica, a computação e o armazenamento em nuvem e a curadoria de dados digitais são fundamentais para continuarem a corresponder com qualidade às exigências da sua comunidade institucional e ao novo perfil dos seus utilizadores que vivem, igualmente, um processo de transformação incessante, imposto pelo ritmo dos avanços tecnológicos no seu quotidiano de trabalho de vida, investigação e estudo.

Palavras-chave: Futuro das bibliotecas; Curadoria digital; Web semântica; Nuvem; Competências dos bibliotecários.

Abstract

The 3rd millennium is the information society era and the age of change, where everything is technological in people's everyday lives. Libraries, and health libraries specially, working everyday with sensitive scientific information, increasingly require to achieve, the most precise, credible and updated information and for that reason, need to adapt their services and collections management to new technology requirements constantly, in order to deliver specialized services and high quality information that best fit in the profile of their users. Currently, libraries are both knowledge producers and spreaders: they are producing contents, collaborating on networks, providing applications for mobile devices, providing digital and remote services and preserving digital data that can be reused in future to create new knowledge. In this environment, it is mandatory that librarians develop new skills to be possible to offer new services and contents, learn new ways to manage their collections (now, also digital collections) and retrieve information, more quickly and efficiently. This communication, gives emphasis to new trends that libraries should follow in the future (that is already present) and where some relevant concepts like semantic web, computing and cloud storage and digital curation, are important key performers, in order to continue delivering the high level of quality required by their institutional community and by the specific profile of their users, who also experience themselves a process of constant transformation, imposed by the technological advances in their daily work life, research and study.

Keywords: Library future; Digital curation; Semantic web; Cloud; Librarian skills.

Introdução

As Bibliotecas de saúde, sobretudo as que estão ligadas à universidade, passaram por diversas mudanças na última década, às quais se tiveram que adaptar, planeando a oferta de novos serviços mais adequados às novas exigências. Entre as mudanças conjeturais destacam-se:

- a) A introdução do processo de Bolonha no ensino superior que tornou o ensino-aprendizagem mais colaborativo e dinâmico, privilegiando o reforço das atividades de investigação científica na comunidade académica e valorizando a literacia da informação como ferramenta de apoio à investigação;
- b) A globalização da informação e a mobilidade de alunos e investigadores;
- c) A penetração da Internet, tecnologias móveis e ferramentas sociais no quotidiano dos indivíduos transformou-os em produtores de informação e conteúdos, numa sociedade da informação onde atualmente quase tudo é móvel (os recursos, as pessoas, as comunicações, os serviços, o conhecimento) e permitiu às bibliotecas desenvolverem novos ambientes digitais e serviços baseados nestas novas tecnologias emergentes;
- d) A existência de uma maior democratização e rapidez de acesso à informação científica pela partilha em redes colaborativas na *web*, nos repositórios e dinamizado através do Open Access.

Assim, para além de oferecerem novos serviços baseados nas novas tecnologias, como o *eLearning*, a produção de conteúdos para dispositivos móveis, os estudos bibliométricos, os serviços de referência digital e a interação à distância com os utilizadores, utilizando as ferramentas web 2.0, a par de outros serviços específicos como, por exemplo, os serviços de informação ao paciente, as Bibliotecas de saúde terão que percorrer novos caminhos para superarem estes novos desafios e continuarem a ser uma marca de excelência, a que habituaram os seus utilizadores.

A gestão e preservação das coleções digitais tornam-se, neste contexto, uma prioridade, beneficiando das inovações tecnológicas que permitem, através da *web* semântica e das ontologias, uma melhor organização e recuperação da informação digital.

Ao introduzirmos processos tecnológicos no processo de tratamento documental, estamos a imprimir eficácia e eficiência à gestão das coleções e a fazer uma gestão mais racional do tempo empregue pelas equipas nestas rotinas, que poderá ser aproveitado de outra forma, quer na produção de outros produtos e conteúdos quer na conceção e oferta de novos serviços em prol do utilizador como, por exemplo, o caso do envolvimento em equipas de projetos multidisciplinares, vocacionadas para a indexação automática dos documentos digitais. É neste contexto que a área da inteligência artificial tem vindo a assumir relevância relativamente à organização e representação do conhecimento, causando impacto no seio das bibliotecas pela utilização de sistemas informáticos com componentes de inteligência artificial, no âmbito da representação dos documentos, na melhoria das estratégias de pesquisa e na formulação e reconhecimento das intenções subjacentes às questões efetuadas pelos utilizadores.

Desenvolver estudos e projetos no âmbito da *web* semântica é, não só uma perspetiva futura, como uma necessidade de desenvolvimento para as bibliotecas que se tem vindo a evidenciar. O caminho ainda é longo, pois ainda há bastantes dificuldades, sobretudo ao nível do desenvolvimento das competências dos bibliotecários neste domínio, como também a dificuldade de integrar equipas multidisciplinares com informáticos para o desenvolvimento deste tipo de estudos e aplicação dos mesmos a novos projetos e realidades.

A utilização da “nuvem” nas bibliotecas, quer para o armazenamento de dados e documentos quer para o desenvolvimento de serviços e aplicações, traz vantagens consideráveis em termos de custos com a aquisição de *hardware/software*, mobilidade e portabilidade da informação e participação em processos de investigação colaborativos e à distância, existindo, porém, algumas desvantagens que poderão ser, no entanto, contornáveis a curto prazo, como é o caso da segurança dos dados e a questão da ausência de legislação neste âmbito.

A questão da preservação de documentos, sobretudo a preservação digital, é uma temática que nas últimas décadas tem vindo a preocupar as bibliotecas. Os problemas são diversos e têm a ver sobretudo com a rápida obsolescência das tecnologias envolvidas na preservação dos documentos devido aos rápidos avanços tecnológicos; a durabilidade dos suportes que pode colocar a médio prazo os documentos digitais em situações de fragilidade e as dificuldades de

gestão dos próprios metadados com qualidade suficiente para permitir uma boa recuperação a *posteriori*.

Também no seio da comunidade médica, académica e nos laboratórios de investigação, a produção, gestão e manuseamento de uma larga dimensão de dados científicos numa base diária colocam a questão do modo como se preservarão esses dados e transformarão nos próximos anos, assumindo-se a curadoria de dados como um serviço extremamente valioso e com os olhos postos no futuro, onde as bibliotecas deverão intervir.

Todas estas mudanças sentidas nas Bibliotecas pela aplicação das tecnologias e pelo eclodir de novas coleções de âmbito digital, a par de novos procedimentos na gestão e preservação documental, produzirão necessariamente mudanças nas competências dos seus bibliotecários, que terão que se adaptar aos novos tempos, novos procedimentos e com pro atividade, oferecer serviços centrados nas necessidades dos seus utilizadores, que também se vêm confrontados no seu dia-a-dia com este novo ambiente digital.

Esta comunicação tem por base a investigação realizada no âmbito do Curso de Doutoramento em Ciências da Informação da Universidade de Évora, do qual se selecionaram algumas áreas que pareceram pertinentes mencionar, no âmbito dos novos caminhos que as Bibliotecas terão que trilhar num futuro próximo, sobretudo as Bibliotecas de saúde, pela constante evolução e transformação dos dados científicos manuseados pelos seus investigadores. Ainda neste âmbito, pretendemos apresentar alguns dados de um inquérito realizado a Bibliotecas académicas portuguesas e europeias e a utilizadores universitários portugueses, tendo por base o euro-referencial I-D do *European Council of Information Association* e que evidenciam quais as competências atuais e a desenvolver futuramente pelos bibliotecários para corresponderem aos novos desafios impostos pelo ambiente tecnológico em que se movem.

Abordaremos, nos capítulos seguintes, aqueles que nos parecem ser os três caminhos mais evidentes a trilhar pelos bibliotecários, nomeadamente os da saúde, num futuro que é já o presente e onde as tecnologias estarão lado a lado com estes profissionais, na gestão, organização, preservação e difusão dos seus serviços e coleções.

Gestão e preservação de coleções – ontologias e web semântica

A gestão e preservação de coleções têm sido a missão primordial das Bibliotecas através dos tempos, no sentido de transmitir o património documental às gerações vindouras e encontrar técnicas cada vez mais sofisticadas de tratar a informação, com vista à sua futura recuperação, de uma forma mais rápida e mais assertiva. Porém, o crescimento desordenado e caótico da informação veio dificultar essa tarefa. Com o advento das novas tecnologias e a sua aplicação às Bibliotecas, este processo saiu reforçado por via da informatização dos catálogos, simplificando a vida ao bibliotecário numa série de tarefas e rotinas, até aí manuais, que passaram gradualmente a integrar os sistemas de gestão documental automatizados e que se tornaram cada vez mais sofisticados com a adoção de práticas da área da inteligência artificial, como as ontologias e a web semântica. Ao introduzirmos processos tecnológicos nestas tarefas, estamos a imprimir eficácia e eficiência à gestão das coleções e a fazer uma gestão mais racional do tempo empregue pelas equipas nestas rotinas, que poderá ser aproveitado de outra forma, quer na produção de outros produtos e conteúdos quer na conceção e oferta de novos serviços em prol do utilizador como, por exemplo, o caso do envolvimento em equipas de projeto multidisciplinar vocacionadas para a indexação automática dos documentos digitais.

É neste contexto que a área da inteligência artificial tem vindo a assumir relevância, relativamente à organização e representação do conhecimento, causando impacto no seio das Bibliotecas pela utilização de sistemas informáticos com componentes de inteligência artificial¹, no âmbito da representação dos documentos (catalogação, classificação, thesauri, indexação e catálogos em linha), na melhoria das estratégias de pesquisa e na formulação das questões geradas pelos utilizadores no sentido do apoio à tomada de decisão. Assim, uma das principais vantagens da utilização da inteligência artificial está relacionada com o aumento da *performance* e melhoria do desempenho dos utilizadores, uma vez que traduz a linguagem natural da questão colocada pelos utilizadores, ajudando-os através do processamento dessa linguagem para um sistema de indexação compreendido pelo computador, de modo a retornar ao utilizador a informação pertinente de auxílio à tomada de decisão, num curto espaço de tempo e com a qualidade e pertinência desejadas. Este sistema de indexação compreendido pelo computador designa-se de

analisador semântico², que reconhece e processa palavras e regras gramaticais, decodificando-as para que o sistema informático possa iniciar um processo de raciocínio que conduza ao resultado da pesquisa e à satisfação da pergunta formulada pelo utilizador. O advento da Internet e a sua utilização massiva foi o móbil para que nos últimos anos se introduzisse neste domínio o conceito de ontologia e *web* semântica e que a sua aplicação viesse a desenvolver extraordinariamente os processos de organização e recuperação da informação em ambiente digital.

Nos processos da inteligência artificial, as ontologias definem o modo como se relacionam os conceitos num domínio específico do conhecimento, sendo que a análise ontológica implica também a interpretação semântica de cada um dos termos contidos nessa ontologia.

Carlan³ apresenta a diversas partes que compõem uma ontologia:

- a) **Classes ou conceitos** – Que descrevem o conceito do domínio e podem dividir-se em classes e subclasses. Ex: vinho é a classe e vinho branco a subclasse.
- b) **Atributos** – São as propriedades das classes. Ex: cor, sabor, corpo.
- c) **Facetas** – Restrições da propriedade. Ex: uma ou mais variedades de uvas.
- d) **Instâncias** – São as ocorrências dos conceitos e as relações estabelecidas pela ontologia. Ex: a instância Porto Calém é um tipo de vinho da classe vinho do Porto e que interage com os seguintes atributos: cor, cheiro, corpo, etc.

As ontologias podem ser de vários tipos³:

- 1) **Ontologias de alto nível** – Descrevem um evento, um objeto, o espaço e o tempo.
- 2) **Ontologias de tarefas** – Descrevem o vocabulário relacionado com um domínio genérico. Ex: Medicina.
- 3) **Ontologias de aplicação** – Descrevem conceitos dependendo do domínio e de tarefas particulares. Ex: prevenção do tabagismo no âmbito da Medicina Preventiva.

Por conseguinte, a construção de ontologias permite obter maior clareza e assertividade na definição de um vocabulário comum para um grupo de pesquisadores que precisa de partilhar informação num dado domínio.

Como exemplo, podemos focar-nos numa ontologia de doenças, combinada com uma ontologia de terapêuticas e dosagens medicamentosas, incorporadas numa aplicação de medicina baseada na evidência para gerar um sistema de diagnóstico e prognóstico médico.

A construção de ontologias assenta na utilização de linguagens, ou seja, códigos específicos para serem lidos em páginas *web* e entendidos pelos meios eletrónicos no processamento da informação. Todas estas linguagens se baseiam em XML (eXtensible Markup Language), uma linguagem especialmente usada para informação na *web* e que possibilita a criação de *tags* (marcas) para descrever os conteúdos dos documentos digitais, possibilitando a organização da informação e sua posterior recuperação a partir daquele termo marcado. Também utilizado na mesma função está o HTML (Hypertext Markup Language). As linguagens ontológicas, mais conhecidas e utilizadas a nível biblioteconómico, são a OWL (Web Ontology Language) que define ontologias na *web*, permitindo associar a semântica ao conteúdo de dados na *web* e o RDF (Resource Description Framework) que representa o conhecimento através de redes semânticas, representando conceitos e taxonomias de conceitos.

As ontologias permitem, deste modo, representar o conhecimento de modo consensual através de um vocabulário uniforme, eliminando contradições e inconsistências na pesquisa da informação e sobretudo a grande inovação trazida pela utilização das ontologias foi a evolução da *web* 2.0 para a *web* 3.0, ou a designada *web* semântica, trazendo uma nova geração de possibilidades no âmbito do tratamento documental, interação entre pessoas e sistemas informáticos e na recuperação rápida e fiável da informação (Tabela 1).

Tabela 1: Evolução da Internet da Web 1.0 à web 3.0

WEB 1.0	WEB 2.0	WEB 3.0
Internet estática A troca de dados era feita entre máquinas	Internet colaborativa Interação entre máquinas e pessoas que passam a ser produtoras e indexadoras de conteúdos porém as páginas só têm informação lexical.	Internet Inteligente e semântica Representação e interação de dados e aplicações suscetíveis de serem transformados em elementos inteligíveis e agentes inteligentes (intervenção de mecanismos da inteligência artificial)

Pedraza-Jimenez, Codina e Rovira⁴ mencionam que há uma tendência de se confundir as ontologias com as taxonomias e os thesauri. Na perspectiva destes autores, as ontologias podem considerar-se linguagens documentais com diferentes níveis de estrutura, porém, a diferença destas para um thesaurus tradicional é que estão feitas com uma sintaxe compreensível para os computadores e contemplam um conjunto mais amplo de relações que as da classe e subclasses das taxonomias ou de sinonímia ou meronímia^{*} como nos thesauri. Assim, uma ontologia permite uma maior riqueza na sua definição de conceitos e suas relações do que o caso de um thesaurus, podendo ser processadas por aplicações informáticas e suportando processos de inferência automáticos.

No que concerne à pesquisa e recuperação da informação, Quaresma⁵ aponta alguns desafios para que esta represente e comunique a real intenção dos utilizadores, afirmando que os sistemas de recuperação, baseados exclusivamente nas pesquisas em texto livre e orientadas por catálogo, são muito limitados e inadequados. No caso das pesquisas em texto livre, estas são baseadas em reconhecimento de padrões, onde todas as palavras são indexadas para responder às questões e, no caso das pesquisas por catálogo, a informação é estruturada, utilizando-se os thesauri para representar o conhecimento e as suas relações.

Assim, a proposta do autor será o recurso à pesquisa pelo conteúdo semântico dos documentos, através de uma interpretação dos documentos (representação semântica) com recurso a uma ontologia e que represente o conhecimento e as suas relações. Deste modo, a interpretação da interrogação formulada pelo utilizador será feita tendo em consideração a ontologia.

Como se aplica a *web* semântica ao ambiente das Bibliotecas? Que potencialidades retirar destas tecnologias?

Como foi perceptível observar pela análise já efetuada aos conceitos de ontologia e *web* semântica, as vantagens aplicadas aos sistemas de recuperação da informação são imensas, pois permitem uma maior eficiência na recuperação dos documentos com um elevado nível de qualidade.

A questão da interoperabilidade é outra vantagem, uma vez que a publicação de coleções digitais proporciona a criação de *links* entre arquivos e outros recursos na *web*. Assim, a interoperabilidade semântica permite relacionar significados de conceitos entre bases de dados e arquivos diferenciados, de modo que, ao formular uma questão, o sistema consiga interpretar os conceitos e compará-los por grau de similaridade e devolver uma resposta agregada⁶.

Greenberg⁷ apresenta várias semelhanças entre as bibliotecas e a *web* semântica:

- Ambas foram criadas para resolver o problema da gestão do excesso de informação desorganizada.
- Ambas possibilitam a descoberta de conhecimento e o acesso à informação.

^{*} Pedraza-Jimenez, Codina e Rovira definem “meronímia” como: “una relación semántica no-simétrica entre los significados de dos palabras dentro del mismo campo semántico. Se denomina merónimo a la palabra cuyo significado constituye una parte del significado total de otra palabra, denominada ésta holónimo. Por ejemplo, dedo es merónimo de mano y mano es merónimo de brazo; a su vez, brazo es holónimo de mano y mano es holónimo de dedo. Ejemplos: azul es merónimo de color; doctor es merónimo de oficio⁴,p.577.

- Ambas se baseiam no desenvolvimento de normas nacionais e internacionais standartizadas.
- Ambas cresceram devido ao espírito colaborativo dos seus interlocutores.
- Ambas se tornaram uma parte integrante e fundamental da sociedade da informação.

Assim, a web semântica é fundamental atualmente para o desenvolvimento das coleções das Bibliotecas, sobretudo no campo digital, ajudando a clarificar tipos de formatos e de etiquetas (*tags*) e, na catalogação, a *web* semântica pode ser aplicada aos vocabulários controlados e thesaurus, sistematizando e relacionando todos os dados de proveniências diferentes, uniformizando-os. Neste processo, o recurso às ontologias permite uma maior precisão na definição de conceitos e suas relações e possui uma sintaxe compreensível para os computadores.

Ao nível da circulação de documentos e dos serviços de referência, o uso da *web* semântica permite uma melhor estruturação da informação com vista à recuperação dos documentos pelo utilizador final.

Greenberg⁷ aponta alguns aspetos que interferem com a participação dos bibliotecários no domínio da *web* semântica:

- Barreiras da comunicação, pela existência de diferentes tipos de linguagens informáticas e no domínio da *web* semântica.
- Ausência de aplicações *web* amigáveis para produzir informação digital interoperável com a *web* semântica.
- Excesso de tarefas quotidianas que exigem a atenção do bibliotecário e que o impossibilitam de poder investigar e avançar em novos projetos no domínio da *web* semântica, adquirindo estas novas competências.
- Existência de documentação limitada e difícil de interpretar, por quem não é informático, sobre os processos, planos e políticas de implementação de sistemas baseados na *web* semântica.

Codina⁸ sintetiza o conceito de *web* semântica, dizendo que o que nos vem ensinar a *web* semântica não é nada que nós já não soubéssemos como fazer. O que é realmente novo é a perspectiva de se tentar converter toda a *web* na mais gigantesca base de dados que a humanidade alguma vez sonhou:

Si tomas un conjunto de datos y los etiquetas sistemática y exhaustivamente, tienes lo más parecido a la inteligencia(...) Así que, lo que es (genialmente) nuevo en la *web* semântica es la idea de convertir toda la *Web* es la más gigantesca base de datos que la humanidad pudiera haber soñado jamás^{8,p.15}.

Resta-nos pois concluir que após muitos séculos de história no percurso evolutivo das Bibliotecas, não era este também o sonho e a visão da Biblioteca de Alexandria de reunir num mesmo espaço todo o conhecimento humano e as obras de todos os povos da terra?

Preservação digital e curadoria de dados

Para Breeding⁹, a preservação digital é a definição de procedimentos e implementação de medidas que permitam manter, no futuro, os materiais do presente intactos e acessíveis, através de dados e metadados para recuperação futura com qualidade por novas gerações, graças à utilização de modelos como o OAIS (*Open Archive Information System*) que já tem provas dadas quanto ao assegurar de uma sobrevivência a longo prazo dos documentos digitais.

Assim, a preservação digital é apenas um dos componentes do processo de curadoria digital.

A produção, gestão e manuseamento de uma larga dimensão de dados científicos numa base diária por parte dos investigadores, nomeadamente nas áreas científicas como a da saúde, tornam imprescindível a curadoria de dados como uma competência extremamente importante a desenvolver nas Bibliotecas de saúde, uma vez que este processo, tal como é definido pelo *Digital Curation Centre* (<http://www.dcc.ac.uk>), permite assegurar a sustentabilidade e a validade dos dados científicos para o futuro, de modo a que possam ser acedidos e reutilizados, agrupados ou transformados.

A maior parte dos dados científicos são únicos e irrecuperáveis em caso de perda ou destruição. As Bibliotecas podem assumir aqui um papel preponderante na definição de procedimentos de conservação e preservação, cumprindo a sua missão ancestral como protetora do património documental institucional e, ao mesmo tempo, participando como parceira nos projetos de investigação da instituição, através da elaboração de um programa de preservação institucional que permita poupar tempo, recursos e gerir da forma mais adequada o investimento feito. Sem haver esta preocupação de preservação dos documentos digitais, a instituição poderá ficar fragilizada, vulnerável e dependente de dados obsoletos.

Deste modo, a curadoria de dados integra em si não só a preservação digital dos documentos mas todo um ciclo de vida dos mesmos que se torna fulcral quando falamos da preservação e reutilização futura de dados científicos. Exemplificando, imaginemos que uma equipa de investigadores descobre a cura para o H5N1 (vulgo gripe A) e comunica os resultados através de artigos científicos que publica em revistas da especialidade, porém, não disponibiliza para curadoria, preservação e reutilização os dados científicos e as fórmulas que lhe permitiu alcançar esses resultados. Décadas ou séculos mais tarde, quando uma nova epidemia de características semelhantes se abater sobre a população, os cientistas da época terão que recomeçar tudo de novo para conseguirem chegar (talvez tarde demais) aos mesmos resultados ou evoluírem na sua investigação a partir do que já havia sido feito e conquistado pelos seus antecessores.

Este ciclo de vida da curadoria digital compreende os seguintes passos[†]:

- a) **Conceptualização** – Conceber e planear a criação de objetos digitais, incluindo os métodos de captura de dados e as opções de armazenamento.
- b) **Criação** – Produção dos objetos digitais e respetivos procedimentos administrativos, descritivos, estruturais e técnicos no que diz respeito ao arquivo dos metadados.
- c) **Acesso e uso** – Assegurar que os utilizadores conseguem aceder facilmente aos dados.
- d) **Avaliação e seleção** – Avaliar os objetos digitais para selecionar os que requerem uma curadoria e preservação a longo prazo.
- e) **Supressão** – Destruição segura dos objetos digitais que não foram objeto de seleção para preservação.
- f) **Inclusão** – Transferência dos objetos digitais para um arquivo, um centro de dados ou um repositório digital seguro, elaborando-se normas e procedimentos para a gestão desses dados.
- g) **Ação de preservação** – Providenciar ações de preservação dos objetos digitais.
- h) **Reavaliação** – Voltar a avaliar os objetos digitais que falham alguns dos procedimentos pré-estabelecidos.
- i) **Armazenamento** – Manutenção dos dados guardados segundo as normas de segurança.
- j) **Acesso e reutilização** – Assegurar que os dados armazenados estarão sempre acessíveis para a utilização e reutilização de grupos definidos de utilizadores.
- k) **Transformação** – Criação de novos objetos digitais, por exemplo, através da migração para outro formato.

[†] Ciclo de vida da curadoria digital conforme apresentado no site do Digital Curation Centre em: <http://www.dcc.ac.uk/print/node/515>

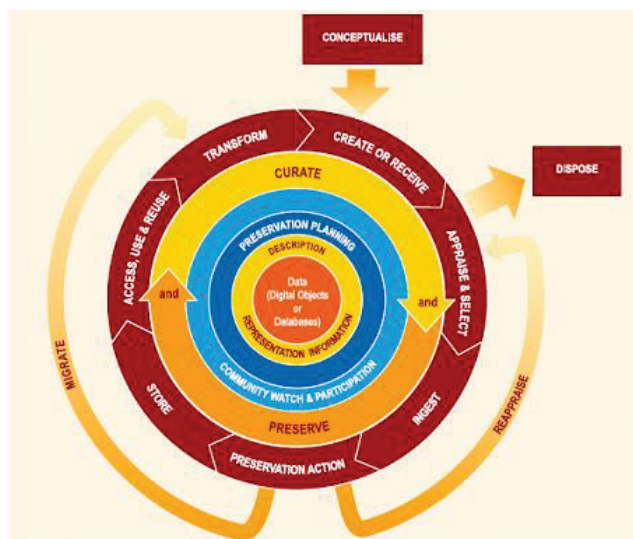


Fig. 1: Modelo do ciclo de vida da Curadoria Digital in: www.dcc.ac.uk

A finalidade da curadoria digital é, em suma, assegurar a sustentabilidade e a validade dos dados científicos para o futuro, de modo a que possam ser acedidos e reutilizados, agrupados ou transformados, sendo que este processo envolve várias fases, desde o planeamento e criação dos dados à descrição e representação da informação através de metadados, à avaliação e seleção de dados para preservação e curadoria a longo prazo, ao arquivo em repositório ou centro de dados, às ações de preservação a longo prazo que mantenham os dados íntegros e fiáveis, ao acesso, utilização e reutilização dos dados e, por fim, à transformação dos dados originais em novos dados.

Serviços “cloud computing” (computação em nuvem)

Convém distinguir, em primeiro lugar, os conceitos de “armazenamento em nuvem” e de “computação em nuvem”.

No “armazenamento em nuvem”, os dados de um utilizador ou organismo são simplesmente guardados em aplicações ou plataformas do fornecedor. Ex: Dropbox.

O conceito “cloud computing” ou “computação em nuvem” surgiu pela primeira vez associado a Eric Schmidt, da empresa Google, numa palestra em 2006¹⁰.

Este modelo funciona como uma prestação de serviços, em que as aplicações/sistemas estão alojados remotamente num servidor de uma empresa externa prestadora de serviços e são acedidos pelo utilizador através do seu PC ou dispositivo móvel que funciona apenas como conexão.

Pode ser representado através do esquema da Figura 2.



Fig.2: Computação em “Nuvem” (20)

Segundo Fenilli e Marchi¹¹, o “cloud computing” assenta em três sistemas:

- a) **Software como um serviço (SaaS):** O *software* e atualizações são disponibilizados via Internet. Ex: Facebook.
- b) **Plataforma como um serviço (PaaS):** O provedor do serviço disponibiliza o sistema operacional, linguagem de programação e o ambiente de desenvolvimento para o utilizador desenvolver as suas próprias aplicações. Ex: Wordpress.
- c) **Infraestrutura como um serviço (IaaS):** O fornecedor que fornece a infraestrutura de armazenamento e processamento e o utilizador cria ambientes de aplicação e gere os sistemas operacionais. Ex: Koha.

Existem ainda, quatro tipos de nuvens que determinam a forma como os serviços são implementados, de um modo aberto ou mais restrito¹²:

a) Nuvem Privada

Neste modelo, a infraestrutura de nuvem é utilizada exclusivamente por uma organização ou empresa, sendo esta nuvem local ou remota e administrada pela própria empresa. Possui políticas de acesso a estes serviços, de autorização e de autenticação definidas pelas empresas.

b) Nuvem Pública

Neste modelo, a infraestrutura de nuvens é de acesso aberto, disponível para todo o tipo de público e pode ser acedido de qualquer parte sem restrições de acesso.

c) Nuvem Comunidade

Neste caso existe uma partilha de uma nuvem por diversas entidades que, pertencentes ao mesmo grupo, partilham interesses e decidem o nível de acesso e segurança, definindo políticas próprias.

d) Nuvem Híbrida

Este modelo agrega várias tipologias de nuvem: pública, privada ou comunitária e que permanecem como entidades únicas, ligadas por uma tecnologia padronizada ou proprietária que permite a portabilidade de dados e aplicações.

Relativamente às vantagens e inconvenientes deste tipo de serviços podemos constatar, pela análise da Tabela 2, que são as seguintes:

Tabela 2: Vantagens e inconvenientes dos serviços de computação e armazenamento em nuvem

Vantagens
Custos com Software/hardware e manutenção são suportados pelo fornecedor do serviço
Maior mobilidade e portabilidade no acesso à informação disponível e acessível de qualquer parte através de dispositivos móveis.
Os dados científicos podem ser mais facilmente partilhados e trabalhados por grupos de investigadores independentemente da sua localização reduzindo-se o custo com viagens, comunicações e tráfego de email.
Ligação da biblioteca a redes mundiais de informação, grupos de utilizadores redes de bibliotecas para realização de projetos conjuntos.
Desvantagens
Segurança dos dados alojados no fornecedor de serviços.
Ausência de legislação relativa à partilha e utilização de dados científicos e seu manuseamento por grupos / indivíduos oriundos de vários países com políticas de copyright diferentes.

De que maneira podem os serviços de “*cloud computing*” ser aplicados às Bibliotecas?

Peters¹³ vem relembrar que quem tem usado a *web 2.0* nos últimos anos (wikipedia, flickr e twitter) já está a utilizar ambientes de “*cloud computing*”, uma vez que este tipo de aplicações está hospedado em centros de dados que, por sua vez, estão na “*cloud*”.

Goldner¹⁴ afirma que as vantagens de utilizar os serviços de “computação em nuvem” nas Bibliotecas advêm de três fatores:

a) Tecnologias

Com a utilização massiva de dispositivos móveis por parte dos utilizadores, o uso destes serviços em nuvem possibilita um acesso mais rápido e cómodo à informação, onde quer que o utilizador se encontre, garantindo a mobilidade e portabilidade da informação.

Também a disponibilização de *hardware* e *software* por parte dos fornecedores garante uma maior flexibilidade e facilidade na utilização destes serviços, a mais baixos custos para as Bibliotecas, que não terão igualmente que aumentar a capacidade dos seus servidores, verificando-se, assim, uma evolução para um sistema mais dinâmico, interoperável e partilhável com outro tipo de sistemas e de serviços existentes na comunidade institucional.

b) Dados

Através da “*cloud*”, os dados científicos podem ser facilmente partilhados entre grupos e comunidades de investigação restritos, disseminados por vários países, reduzindo-se assim o número de comunicações, viagens e de tráfego de *email*. Desaparece também a necessidade de armazenamento local dos dados e *backups* sucessivos.

Também a disseminação de informação de domínio público pode facilitar a divulgação feita pela Biblioteca e chegar mais facilmente a todos os potenciais interessados.

c) Colaboração em comunidade

É inegável que a grande vantagem da utilização da “*cloud*” é sobretudo a facilidade como se partilha informação, aumentando a facilidade de colaboração interna e externa à comunidade académica, ligando a Biblioteca ao mundo e a redes diversificadas de utilizadores e de Bibliotecas congêneres para a realização de projetos conjuntos. Os serviços “*cloud computing*” dão às Bibliotecas o poder de criar grandes redes *online* de informação onde se gera e partilha conhecimento, demonstrando uma vez mais o seu valor.

Wallis¹⁵ remete-nos para o desenvolvimento dos serviços de catalogação e de gestão de coleções, baseados na aplicação de tecnologias emergentes como o “*cloud computing*” e cuja novidade mais recente foi a inclusão da OCLC (*Online Computer Library Center*) nos serviços “nuvem” através do WorldCat.

Competências a desenvolver pelos bibliotecários para se adaptarem a serviços baseados em novas tecnologias emergentes

No âmbito do curso de Doutoramento em Ciências da Informação da Universidade de Évora, a autora realizou inquéritos a Bibliotecas portuguesas e europeias e a utilizadores universitários portugueses num total de 47 Bibliotecas académicas e 136 utilizadores. No caso português, 33% das Bibliotecas eram da saúde e no caso europeu 82% eram Bibliotecas da saúde. Os utilizadores eram sobretudo alunos de licenciatura (44%) e de mestrado (41%).

No inquérito era pedido que, com base no Euroreferencial I-D (e para os grupos do Euroreferencial: *Informação*, *Tecnologias*, *Comunicação* e *Gestão*), indicassem as competências dos bibliotecários ainda a desenvolver para fazerem face às exigências e transformações impostas pela nova sociedade da informação cada vez mais tecnológica. Os resultados foram os seguintes.

No grupo *Informação* verificamos uma preocupação em todos os grupos inquiridos com a conceção e desenvolvimento de novos produtos e serviços decorrentes da utilização das

tecnologias nas Bibliotecas cada vez mais interativas e multimedia e onde os utilizadores recorrem, solicitando este tipo de apoio mais especializado. Também a preocupação com o direito de autor foi bastante evidente.

Tabela 3: Competências a desenvolver pelos bibliotecários – Grupo do Euro-referencial I-D – Informação

Bib. Portuguesas	Bib. Europeias	Utilizadores
Conceção e desenvolvimento de Produtos e Serviços	Conceção e desenvolvimento de Produtos e Serviços	Conceção e desenvolvimentos de Produtos e Serviços
Aplicação do direito de Informação	Aplicação do direito de Informação	Enriquecimento das coleções e fundos

No âmbito das competências tecnológicas a desenvolver (grupo *Tecnologias*) e em sintonia com o que foi respondido pelos inquiridos no grupo *Informação* (ver Tabela 3), o desenvolvimento informático de novas aplicações é o que mais se evidencia, nomeadamente em casos específicos, como na área da saúde, em que se desenvolvem aplicações específicas sobre estudos de casos clínicos para dispositivos móveis (Tabela 4).

Tabela 4: Competências a desenvolver pelos bibliotecários – Grupo do Euro-referencial I-D - Tecnologias

Bib. Portuguesas	Bib. Europeias	Utilizadores
Desenvolvimento informático de aplicações	Desenvolvimento informático de aplicações	Desenvolvimento informático de aplicações (para dispositivos móveis)
Concepção informática de sistemas de informação documental	Tecnologias Web	Publicação e edição

No Grupo *Comunicação*, todos os inquiridos são unânimes em salientar as principais competências a desenvolver: comunicação audiovisual, prática de uma língua estrangeira e literacia/comunicação informática (Tabela 5).

Tabela 5: Competências a desenvolver pelos bibliotecários – Grupo do Euro-referencial I-D - Comunicação

Bib. Portuguesas	Bib. Europeias	Utilizadores
Comunicação audiovisual	Comunicação audiovisual	Comunicação audiovisual
Prática de uma língua estrangeira	Prática de uma língua estrangeira	Prática de uma língua estrangeira
Literacia/comunicação informática	Literacia/comunicação informática	

Ao nível do grupo *Gestão*, sobressaem as competências de marketing sobretudo o marketing digital (utilizando as redes sociais) (Tabela 6).

Tabela 6: Competências a desenvolver pelos bibliotecários – Grupo do Euro-referencial I-D - Gestão

Bib. Portuguesas	Bib. Europeias	Utilizadores
Venda e difusão		Venda e difusão
Marketing	Marketing	Marketing
	Gestão de Projeto e planificação	Formação e ações pedagógicas

Os bibliotecários são, desta forma, parceiros na sua comunidade, no âmbito da investigação, da literacia da informação e informática, na produção de conteúdos e novos serviços baseados nas tecnologias e no ensino-aprendizagem cada vez mais interativo e colaborativo a desenvolver-se num ambiente híbrido em que as Bibliotecas são agora verdadeiros laboratórios de aprendizagem e de experimentação para utilizadores e bibliotecários.

Conclusão

As Bibliotecas de saúde estão perante um mundo novo que o ambiente digital veio trazer às suas práticas quotidianas. Ao longo das últimas décadas, as Bibliotecas souberam corresponder de imediato a todas as transformações, concebendo novos serviços e conteúdos de acordo com as necessidades e novo perfil dos seus utilizadores. Chegou a vez de olhar para as suas coleções que, nos últimos anos, também sofreram transformações e são atualmente híbridas com enfoque no digital. Chegou também o tempo em que as Bibliotecas participam cada vez mais na visão e estratégia de desenvolvimento da sua comunidade e em que o apoio à investigação científica é crucial para a sustentabilidade da própria instituição. Assim, os processos de curadoria dos dados científicos e a sua ligação às publicações científicas dos investigadores são um passo em frente para assegurar a continuidade, a qualidade e o desenvolvimento da investigação no futuro por outros investigadores. E preservar o passado assegurando a continuidade e reutilização no futuro é, e sempre foi, a missão primordial das Bibliotecas ao longo da história.

Atualmente, os bibliotecários possuem competências de gestão, de liderança, financeiras, formativas, pedagógicas, tecnológicas e formação académica de nível superior, estando habilitados, como qualquer outro gestor, para coordenar as áreas onde possuem *know how*, atestado por vários anos de formação e experiência profissional em campo. Devem, porém, estar atentos às tendências conjunturais, ao perfil dos seus clientes/utilizadores e saber como implementar serviços no âmbito da *web 2.0*, *web 3.0* e muito proximamente no âmbito da *web 4.0* (serviços de âmbito interativo, tridimensional e nas áreas da inteligência artificial).

Coleções híbridas, objetos digitais e novos serviços baseados em tecnologias são igualmente sinónimo de novos desafios, não só da forma como se poderá potenciar e tornar mais rápida e eficiente a recuperação da informação, sempre que esta é necessária aos nossos clientes/utilizadores, como a questão da preservação das coleções, nomeadamente o modo como se processará a preservação dos documentos digitais e os dados científicos.

Assim, emergem duas áreas que, em nosso entender, virão a tornar-se fundamentais a curto prazo no âmbito da gestão de bibliotecas, que deverão ser objeto de novos estudos e que devem ser assumidas dentro da área de competências profissionais da biblioteconomia. São elas o desenvolvimento da *web semântica* e a preservação digital.

Desenvolver estudos e projetos, no âmbito da *web semântica*, é não só uma perspetiva futura, como uma necessidade essencial, tendo em consideração o desenvolvimento do modelo de Biblioteca académica proposto neste estudo. O caminho ainda é longo, pois ainda há bastantes dificuldades, sobretudo ao nível do desenvolvimento das competências dos bibliotecários neste domínio, como também a dificuldade de integrar equipas multidisciplinares com informáticos para desenvolvimento deste tipo de estudos e aplicação dos mesmos a novos projetos e realidades.

É tempo de mudança. É tempo de transformação. E as tecnologias são ventos que sopram a favor da ciência, da investigação do ensino e das Bibliotecas, esses centros de excelência onde circula o conhecimento de alto nível.

Referências bibliográficas

1. Viana CL. O impacto das inteligências artificiais na formação dos bibliotecários e cientistas da informação: revisão de literatura. *Ciênc Inf.* 1990 [cited 2013 Apr 27];19(1):72-8. Available from: <http://eprints.rclis.org/7913/>
2. Mendes RD. Inteligência artificial: sistemas especialistas no gerenciamento da informação. *Ciênc Inf.* 1997 [cited 2011 Jun 6];26(1):39-45. Available from: http://www.scielo.br/scielo.php?pid=S0100-19651997000100006&script=sci_arttext
3. Carlan E. Ontologia e web semântica. Brasília: Universidade de Brasília; 2006 [cited 2013 May 30]. Available from: http://www.researchgate.net/publication/28805219_Ontologia_e_web_semntica
4. Pedraza-Jiménez R, Codina L, Rovira C. Web semántica y ontologías en el procesamiento de la información documental. *Profes Inf.* 2007 [cited 2012 Jan 12];16(6):569-78. Available from: <http://www.lluiscodina.com/webSemanticaOntologias2007.pdf>
5. Quaresma P. (2006) Bibliotecas digitais: novos desafios para os sistemas de recuperação da informação. In 3ª Conferência do Cenáculo, Universidade de Évora, 14-15 de Novembro de 2006 [cited 2013 May 17]. Available from: http://www.di.uevora.pt/~pq/RI_cenaculo.pdf
6. Hollink L, Isaac A, Malaisé V, Schreiber G. Semantic web opportunities for digital libraries. In Proceedings of 32nd Library Systems Seminar of the European Library Automation Group (ELAG). 2008 [cited em 2012 Jun 6]. Available from: http://library.wur.nl/elag2008/presentations/laura_hollink.pdf
7. Greenberg J. Advancing the semantic web via library functions. *Cataloging Classif Quar.* 2007 [cited 2012 Jun 6];43(3-4):203-25. Available from: http://www.tandfonline.com/doi/pdf/10.1300/J104v43n03_11
8. Codina L. Internet invisible y web semántica: ¿el futuro de los sistemas de información en línea?. *Tradumática: traducció i technologies de la informació i la comunicació.* 2003 [cited 2013 May 30];2(Nov):16 p. Available from: <http://www.fti.uab.es/tradumatica/revista/num2/articles/06/06central.htm>
9. Breeding M. Digital archiving in the age of cloud computing. *Comput Libr.* 2013 [cited 2013 May 7];33(2):22-6. Available from: <http://web.ebscohost.com/ehost/pdfviewer/pdfviewer?sid=77e9975c-544f-4f49-9e68-3e447441d27e%40sessionmgr110&vid=4&hid=113>
10. Taurion C. Cloud computing: computação em nuvem: transformando o mundo da tecnologia da informação. Rio de Janeiro: Brasport; 2009 [cited 2013 May 27]. Available from: [http://books.google.pt/books?id=mvir2X-A2mcC&pg=PR4&lpg=PR4&dq=TAURION,+Cezar+\(2009\).+Cloud+Computing:++Computa%C3%A7%C3%A3o+em+Nuvem:+Transformando+o+mundo&source=bl&ots=C8Li1yTNUz&sig=SJUfAr6q7gwAyWs5c0qSNPpYth4&hl=pt-PT&sa=X&ei=06moUdifD-il7Abq24CYBw&ved=0CDEO6AEwAA#v=onepage&q=TAURION%20Cezar%20\(2009\).%20Cloud%20Computing%3A%20Computa%C3%A7%C3%A3o%20em%20Nuvem%3A%20Transformando%20o%20mundo&f=false](http://books.google.pt/books?id=mvir2X-A2mcC&pg=PR4&lpg=PR4&dq=TAURION,+Cezar+(2009).+Cloud+Computing:++Computa%C3%A7%C3%A3o+em+Nuvem:+Transformando+o+mundo&source=bl&ots=C8Li1yTNUz&sig=SJUfAr6q7gwAyWs5c0qSNPpYth4&hl=pt-PT&sa=X&ei=06moUdifD-il7Abq24CYBw&ved=0CDEO6AEwAA#v=onepage&q=TAURION%20Cezar%20(2009).%20Cloud%20Computing%3A%20Computa%C3%A7%C3%A3o%20em%20Nuvem%3A%20Transformando%20o%20mundo&f=false)
11. Fennilli AT, Marchi KR. Computação em nuvem: um futuro presente. Paraná: Universidade Paranaense; 2011 [cited 2013 May 12]. Available from: <http://antigo.unipar.br/~seinpar/artigos/Andressa-Fenilli.pdf>
12. Sousa FR, Moreira LO, Machado JC. Computação em nuvem: conceitos, tecnologias, aplicações e desafios. Ceará: Universidade Federal do Ceará; 2009 [cited 2013 Aug 15]. Available from: http://files.0fx66.com/paper/Computacao_Nuvem.pdf
13. Peters C. What is cloud computing and how will it affect libraries? *Techsoup for Libraries.* 2010 [cited 2013 May 15]. Available from: <http://www.techsoupforlibraries.org/blog/what-is-cloud-computing-and-how-will-it-affect-libraries>
14. Goldner MR. Winds of change: libraries and cloud computing. OHIO: OCLC; 2010 [cited 2012 Mar 17]. Available from: <http://www.oclc.org/content/dam/oclc/events/2011/files/IFLA-winds-of-change-paper.pdf>

15. Wallis R. Linking the clouds. SCONUL FOCUS. 2010 [cited 2013 May 12];50:12-3. Available from: http://www.sconul.ac.uk/sites/default/files/documents/4_4.pdf

Notas biográficas

Paula SARAIVA. Doutorada em Ciências da Informação e da Documentação pela Universidade de Évora, é Chefe de Divisão na Biblioteca do ISCSP. De 2000 a 2010 foi, na Faculdade de Medicina de Lisboa, coordenadora do núcleo de difusão da informação e biblioteca digital e, em 2008, foi Chefe de Divisão do Instituto de Formação Avançada. Foi ainda de 2005-2007 membro do Conselho Diretivo e, de 2005 a 2009, Docente Livre em Literacia da Informação.

Co-Chair da Comissão Científica da 12ª Conferência da EAHIL, (Estoril, 2010), integrou ainda as Comissões Científicas da 13ª Conferência da EAHIL (Bruxelas, 2012), das X Jornadas da APDIS, (Lisboa, 2012) e *do Workshop* da EAHIL (Estocolmo, 2013).

É membro da APDIS (vogal do Conselho Fiscal de 2003 a 2009), da BAD (*vogal do conselho fiscal de 2014-2016*) e da EAHIL desde 2000.

Paulo QUARESMA. É Professor Associado no Departamento de Informática da Universidade de Évora. Foi Diretor da Escola de Ciências e Tecnologia da referida Universidade de 2009 a 2013 e é membro de Comissões de Curso de Informática e de Ciências de Informação e Documentação. Nos últimos anos orientou diversos trabalhos de Doutoramento e de Mestrado nestas áreas, tendo publicado várias dezenas de artigos científicos em revistas e conferências internacionais.