

Engenharia de Documentos e suas relações com a Arquivologia, a Diplomática Digital e a Ciência da Informação

Eduardo Watanabe

Mestre; Universidade de Brasília, Brasília, DF, Brasil;
edw.sfs@gmail.com

Renato Tarciso Barbosa de Sousa

Doutor; Universidade de Brasília, Brasília, DF, Brasil;
renasou@unb.br

Resumo: A Engenharia de Documentos é uma disciplina da Ciência da Computação que investiga os documentos em meio eletrônico. O objetivo deste artigo consiste em comparar e articular os diferentes enfoques da Engenharia de Documentos, Arquivologia, Diplomática Digital e Ciência da Informação sobre documentos e informações. A metodologia utilizada é a pesquisa documental da literatura que inclui a Diplomática Digital, a Tipologia Documental e as funções arquivísticas de Criação e Descrição. Como resultado foi identificado a proximidade entre os conceitos sobre documentos utilizados pelas diferentes áreas de conhecimento e a concepção do documento a partir de processos de negócio ou conjuntos documentais. As principais diferenças residiram na falta de equivalência dos conceitos de espécie e tipos documentais utilizados pela Arquivologia; na lacuna do tratamento de questões de autenticidade de documentos nato-digitais objeto da Diplomática Digital.

Palavras-chave: Engenharia de Documentos. Arquivologia. Ciência da Informação. Diplomática Digital. Tipologia Documental.

1 Introdução

A Engenharia de Documentos é uma disciplina da Ciência da Computação que investiga os sistemas eletrônicos de documentos em qualquer forma e em todas as mídias a partir de princípios, ferramentas e processos que aperfeiçoem a criação, o gerenciamento e a manutenção de documentos (MUNSON, 2001). Não obstante a Arquivologia e a Ciência da Informação também tratem dos documentos como objeto de estudo direta ou indiretamente, há uma lacuna na literatura que as relacione à Engenharia de Documentos.

Considerando a importância crescente da tecnologia da informação, formulamos a seguinte pergunta: como a Engenharia de Documentos pode contribuir com a Arquivologia e a Ciência da Informação? Nesse sentido, o objetivo geral deste trabalho consiste em comparar e articular os seus diferentes enfoques em relação aos documentos e as informações.

A metodologia a ser utilizada é a pesquisa documental combinada com a análise qualitativa dos documentos. A seleção de trabalhos, entre 1900 e abril de 2019, foi iniciada pela busca textual com os operadores: **engenharia de documentos**, *document engineering*, *ingénierie documentaire* e *ingeniería documental*, nas seguintes bases: Base de Dados em Ciência da Informação (BRAPCI), Portal de Periódicos da CAPES, Catálogo de Teses e Dissertações da CAPES, Biblioteca Digital Brasileira de Teses e Dissertações, *Bielefeld Academic Search Engine* (BASE), *Directory of Open Access Journals* (DOAJ), *Networked Digital Library of Theses and Dissertations* (NDLTD). Foram identificados 526 trabalhos no Portal da CAPES de 1969 a 2019, 120 na BASE de 1990 a 2019, 1.785 na DOAJ de 2000 a 2019, 488 na NDLTD de 1987 a 2019 e nenhum trabalho na BRAPCI e Biblioteca de Teses e Dissertações. No processo de pesquisa foram muito importantes as consultas às referências citadas nos trabalhos.

No que se refere à Arquivologia, à Diplomática Digital e à Ciência da Informação, a pesquisa documental foi direcionada para a literatura que aborda assuntos com potencial de interesse para a Engenharia de Documentos por se relacionarem com funções, atividades e modelos de documentos eletrônicos: Diplomática Digital, Tipologia Documental e as funções arquivísticas de Criação e Descrição.

Após a revisão da literatura, foram selecionados e articulados os conceitos que aproximam as diferentes áreas. O presente artigo está organizado em seções sobre a evolução da Engenharia de Documentos, a apresentação de temas relacionados à Arquivologia, à Diplomática Digital e à Ciência da Informação, análise dos resultados e conclusões.

2 Evolução da Engenharia de Documentos

O advento do digital serve ao início da razão computacional, que abre caminho para novos métodos de registro que induzem a constituição de modos inovadores de representação (CROZAT, 2012). Roger Pédaque (2007¹ *apud* CROZAT, 2012) considera que o digital favorece a produção e a transmissão de documentos, o que gerou uma revolução conhecida como redocumentarização, que corresponde ao retorno à antiga documentarização, mas em um novo suporte fornecido pela evolução tecnológica e diante mudanças sociais percebidas nas múltiplas dimensões do documento e nos movimentos de suas mediações.

De acordo com Ayuang (2004² *apud* BACHIMONT, 2007), a ciência busca descobrir conhecimentos que ainda não temos, ao passo que a técnica visa inventar artefatos que ainda não existem. As leis universais são liberadas pela ciência, como o que é válido sempre e em toda parte, o que se repete sem exceção. Já a técnica constrói o que não poderia existir, pois é na contingência que ela se encontra. Só que a técnica também faz parte das possibilidades do Ser na medida em que inventa onde a necessidade da natureza deixa a possibilidade (BACHIMONT, 2007). A técnica então inventa as possibilidades em vez de apenas preenche-las com artefatos, momento em que abre novos horizontes para a ação, construindo o que as leis do Ser não nos permitiram antecipar.

Por sua vez, o *savoirfaire* (*know-how*) gerado pela técnica dá origem à implementação da Engenharia, que contém

[...] objetivos contingentes e utilitários ao exame teórico e reflexivo ao mergulhar no pensamento abstrato e científico idealizado na complexidade do concreto, que os exerce e os faz progredir. (BACHIMONT, 2007, p. 12, tradução nossa).

Assim, a Engenharia não se resume a uma ciência aplicada que se limita a implementar o conhecimento teórico em um contexto prático, pois ela vai além e adentra uma dimensão original do conhecimento humano que excede a postura científica. A evolução da tecnologia permite o construir e o “fazer coisas” que a ciência ainda não entende, de modo a qualificar o papel da técnica também para expandir novos horizontes.

O primeiro estágio dessa evolução é o da engenharia artesanal, em que o mundo é pensado no modelo de uma oficina de trabalhos manuais com o uso improvisado de ferramentas para se ajustar à natureza única dos materiais utilizados e ao produto singular a ser obtido ao final (BACHIMONT, 2007). No segundo estágio, o da engenharia industrial, o mundo passa a ser projetado no modelo da fábrica, a partir de um sistema de produção que se utiliza de materiais mais brutos (*crus*), só que eles já estão pré-formatados para a produção, que é assim marcada pela repetibilidade e padronização. É como engenharia industrial que a Engenharia de Documentos surge, relacionada à escalabilidade das soluções, à reusabilidade e ao intercâmbio de dados e padrões, com processo de validação similar ao processo produtivo industrial físico (PIEZ, 2001).

A origem da Engenharia de Documentos, como disciplina, remonta a uma apresentação feita em 1967 por William Tunnicliffe, presidente da *Graphic Communications Association*, na qual enunciou a necessidade de separação do conteúdo intelectual dos documentos da sua apresentação (GEURTS, 2010). Essa noção de separação do conjunto de dados contidos no documento de sua apresentação é conhecida na literatura da área como o “paradigma da engenharia de documentos”, “modelo de múltiplas entregas de publicação” ou “separação do conteúdo do estilo” (GEURTS, 2010, p. 14, tradução nossa).

Quint, Nanard e André publicaram em 1990 o artigo *Towards Document Engineering*, no qual comparam os métodos e técnicas utilizados na Engenharia de *Softwares* com os usados para edição de documentos eletrônicos:

Quadro 1 - Comparativo entre Engenharia de *Softwares* e Engenharia de Documentos

Característica	Engenharia de <i>Softwares</i>	Engenharia de Documentos
Finalidade	<i>Softwares</i> podem ser considerados um tipo de documento, mas a sua finalidade é muito específica: ser processado por um computador.	Um documento é elaborado com o objetivo de ser lido por humanos, mas que se for também eletrônico deve ser formatado para ser processado por muitas aplicações.
Semântica	O <i>software</i> deve ser representado em um computador e capaz de ser processado por ele.	Um documento estruturado (<i>logical</i> ou <i>abstract structure</i>) pode ser armazenado em um banco de dados, o que: facilita a

		utilização de sistemas de recuperação da informação; permite a sua transformação em diversas formas; viabiliza a sua reorganização; possibilita que diferentes aplicações possam processá-lo. Considerando os diversos usos do documento, sua semântica não é bem definida como a de um <i>software</i> .
Elicitação de requisitos	O programador em geral deve elicitar os requisitos do <i>software</i> durante a fase de especificação, antes de escrever o programa.	O autor do documento muito raramente precisa elicitar requisitos para a sua elaboração.
Estrutura lógica	A representação da estrutura sintática do <i>software</i> consiste em uma estrutura de árvore (hierárquica) que deve ser utilizada por todos os seus editores em regra. A estrutura do <i>software</i> está definida completamente por uma gramática, que deve ser atendida de forma consistente.	Os documentos necessitam de estruturas mais complexas que podem combinar uma estrutura de árvore com outros relacionamentos não-hierárquicos (referências cruzadas e indexações). As características mais valorizadas são as estruturas não-hierárquicas por viabilizar a navegação pelo documento, estabelecer os vínculos entre documentos ou compartilhar partes deles. O documento segue pelo menos duas gramáticas: a da estrutura lógica e a de linguagem natural em que ela está escrita.
Limitações da estrutura lógica	Nos <i>softwares</i> está bem definida a separação entre estrutura sintática, aspectos gráficos e texto livre.	Nos documentos, os elementos estruturais de estrutura lógica, estrutura física e conteúdo não possuem limites bem definidos.
Aspectos gráficos	A estrutura gráfica é geralmente constituída por uma sequência de linhas indentadas com poucas variações de estilo de fonte.	Possuem em parte ou no todo estrutura gráfica mais complexa, com estruturas gráficas bidimensionais. Devido à sua finalidade precípua para leitura humana, a aparência gráfica do documento é muito importante, de modo que muitos escritores consideram primeiro a aparência gráfica antes da sua

		representação abstrata ou lógica.
Processo de escrita	Muitos são projetados de uma forma <i>top-down</i> , com sucessivos refinamentos. Uma técnica recomendada é que o programa seja elaborado em pequenos módulos em lugar de um grande monolito, o que garante maior clareza, reusabilidade, manutenção e compartilhamento.	O autor inicia a escrita de texto desestruturado, apenas para capturar seus pensamentos, sendo que a fase de estruturação só vem depois. A modularização do documento é mais difícil e o encapsulamento não pode ser feito por completo, não obstante eles sejam viáveis.
Perfil do usuário do editor	O usuário do editor de <i>software</i> é um programador que conhece a estrutura sintática do programa e a estrutura hierárquica da sua representação.	O usuário do editor de documento não possui em geral conhecimento sobre a estrutura lógica abstrata do documento, o que requer um estilo diferente de interface.

Fonte: Adaptado de Quint, Nanard e André (1990).

Os documentos possuem peculiaridades em sua produção, como a falta de passos previamente definidos para a sua produção, diferentemente da abordagem sequencial clássica de desenvolvimento de software (QUINT; NANARD; ANDRÉ, 1990). Isso porque a produção dos documentos é livre, não ordenada e, em geral, não é previsível ao lidar com as trocas entre os processos envolvidos de produzir sentença e parágrafos, reunir ideias e dados, revisar e organizar ideias no nível semântico, formatar e estruturar logicamente o documento e administrar a produção dos documentos.

A definição expressa da Engenharia de Documentos demorou a surgir, com a primeira aparição em artigo de Bebo White, da Universidade de Stanford, em 1996, sobre Engenharia de Documentos da *Web*. No quadro abaixo podemos observar a evolução das definições no tempo:

Quadro 2 - Definições de Engenharia de Documentos

Definição	Fonte
Engenharia de Documentos da <i>Web</i> é definida como a aplicação de técnicas de engenharia de software ao design de documentos Web . Ela também incorpora técnicas únicas para o suporte e desenvolvimento no	Bebo White (1996, p. 1, tradução nossa, grifo)

meio hipertexto ou hipermídia que tenha sido resultado de pesquisa e desenvolvimento em sistemas hipertexto.	nosso)
Engenharia de documentos é o desenvolvimento sistemático de apresentações, representações e ferramentas de documentos, que deve basear-se em modelos formais explícitos de documentos e não em representações de documentos.	Anne Brüggemann-Klein e Derick Wood (2000, p. 2, tradução nossa, grifo nosso)
A engenharia de documentos é uma disciplina emergente dentro da ciência da computação que investiga sistemas para documentos em qualquer forma e em todas as mídias. A engenharia de documentos preocupa-se com princípios, ferramentas e processos que melhoram nossa capacidade de criar, gerenciar e manter documentos , assim como a engenharia de software examina os mesmos problemas de <i>software</i> .	Ethan Munson (2001, p. 1, tradução nossa, grifo nosso)
Engenharia de Documentos sintetiza ideias complementares de análise de informação e sistemas, publicação eletrônica, análise de processos de negócio, e informática de negócios. [...] Ela provê os conceitos e os métodos necessários para alinhar a estratégia do negócio com a tecnologia da informação, para preencher a lacuna entre o que se quer fazer e o como fazê-lo .	Robert Glushko e Tim McGrath (2005, p. 5, tradução nossa, grifo nosso)
O objeto da engenharia de documentos é a concepção de sistemas técnicos que permitem otimizar a articulação da manipulação técnica com a interpretação cultural dos documentos.	Stéphane Crozat (2016, p. 6, tradução nossa, grifo nosso)

Fonte: Elaborado pelos autores.

Podemos constatar que a Engenharia de Documentos tem evoluído pela contribuição que cada autor enfatizou em sua época de modo a acumular e sofisticar o conhecimento produzido na área. Como exemplos, mencionamos White (1996) (*design*, hipertexto e hipermídia), Brüggemann-Klein e Wood (2000) (modelos formais explícitos), Munson (2001) (criar, gerenciar e manter documentos), Glushko e McGrath (2005) (estratégia do negócio) e Crozat (2016) (manipulação técnica articulada com interpretação cultural). A intenção aqui foi apenas ilustrar as ideias dos autores a partir da definição por eles apresentada, pois suas obras são muito mais amplas e exigiriam um nível de detalhamento que não compreende a intenção do presente artigo.

Os limites da Engenharia de Documentos parecem estar ainda em consolidação. No *The ACM Symposium on Document Engineering*, o evento mais representativo da área, uma quantidade significativa dos *papers* publicados não faz menção expressa ao termo Engenharia de Documentos, a mesma situação encontrada em outros eventos da área que foram pesquisados.

De todo modo, a Engenharia de Documentos contempla um amplo espectro de abordagens que inclui desde a utilização de análise de estrutura lógica/linguística e reconhecimento de documentos, pesquisa de imagem e multimídia e conversão *raster-vetor* de linhas de arte, mapas e desenhos técnicos, até a geração automática de documentos, documentos para dispositivos móveis e o processamento e interação de documentos *web* (LINS, 2011; SABLATINIG; HASSAN, 2016).

Para fins didáticos, propomos separar a Engenharia de Documentos em dois campos de estudo semelhantes aos que iremos abordar adiante na Diplomática Digital. O primeiro deles lida com os documentos do passado, aqueles que já foram criados e que agora precisam ter as suas informações organizadas para recuperação.

Antona Copoulos, Karatzas, Krawczyk e Wiszniewski (2004) desenvolveram um modelo de estrutura e conteúdo para o ciclo de vida de documentos históricos digitais, que compreende a digitalização, qualificação, segmentação, extração, aceitação e a exploração. Witschurke (2005) elaborou um conjunto de ferramentas para apoiar o processo de digitalização de documentos de arquivo, desde o seu escaneamento até a apresentação do conteúdo a ser recuperado. Já Burkard, Vogeler e Gruner (2008) e Becker, Rauber, Heydegger, Schnasse e Thaller (2008) criaram ferramentas para a marcação XML de documentos medievais digitalizados e seus impactos nas Ciências da História para apoiar a preservação de documentos digitais.

O segundo campo envolve uma abordagem prospectiva para a criação de modelos de documentos nato-digitais a serem utilizados em aplicações no futuro:

Quadro 3 - Pesquisas na Engenharia de Documentos sobre documentos transacionais

Descrição do estudo	Fonte
<p>1. Desenvolver um modelo de documento. 2. Desenvolver representações para os documentos que seguem o modelo. 3. Definir requisitos de funcionalidades e desempenho para operações em documentos no modelo, baseado em aplicações de documentos. 4. Implementação do modelo, que envolve: a) especificar e implementar classes de documentos (no sentido de orientação a objeto, não de definições de tipo de documento) para vários grupos de requisitos de funcionalidade e desempenho. A especificação da classe deve ser com base no modelo, não nas representações de documentos; b) o fornecimento de interfaces para as representações do documento (analisável e não analisável, importação e exportação).</p>	<p>Anne Brüggemann-Klein e Derick Wood (2000)</p>
<p>Sua perspectiva unificada centrada-em-documentos (<i>document-centric</i>) nos ajuda a conceber e entender os novos modelos de negócio baseados em rede viabilizados pela Internet e tecnologias de suporte. A essência da Engenharia de Documentos é analisar e projetar métodos que produzam:</p> <ul style="list-style-type: none"> • a) especificações precisas ou modelos requeridos pelos processos de negócio; • b) regras pelas quais os processos relacionados são coordenados, se entre diferentes firmas para criar serviços combinados ou empreendimentos virtuais ou dentro de uma empresa para agilizar o fluxo de informações entre organizações. <p>Nova disciplina destinada a especificar, projetar, e implementar os documentos que servem como as interfaces para os processos de negócio. [...] modelos que enfatizam os requisitos e padrões de documentos de interoperabilidade da informação.</p>	<p>Robert Glushko e Tim McGrath (2005, tradução nossa)</p>

Fonte: Elaborado pelos autores.

Nessa abordagem prospectiva, o objetivo da Engenharia de Documentos consiste em modelar documentos e serviços que sejam práticos e efetivos (GLUSHKO; McGRATH, 2005). Ela tem raízes interdisciplinares na análise de informações e sistemas, na publicação eletrônica, na análise de processos de negócio e informática de negócios, e o design centrado no usuário.

3 Diplomática, Tipologia Documental, as funções arquivísticas de Classificação, Descrição e Criação e a Diplomática Digital

Um dos objetivos da Arquivologia consiste no acesso à informação, seja ela para subsidiar o processo decisório, para servir como defesa dos direitos do cidadão ou para preservar a memória da sociedade e seus atores. Tudo isso é feito por meio da organização de dados contidos nos arquivos com vistas à informação.

A Diplomática é a disciplina que estuda o documento único, a unidade arquivística elementar, para analisá-lo em seus aspectos formais com o propósito de definir a natureza jurídica dos atos tanto na sua formação como nos seus efeitos (CARUCCI, 1987³ *apud* TOGNOLI, 2013). Ela estuda a espécie documental, assim entendida como o modelo prescrito pelas normas jurídicas, com redação que segue um determinado padrão semântico e formatado de modo a ter validade e credibilidade (BELLOTTO, 2002). A Diplomática e, mais propriamente, a Diplomática Contemporânea ou Tipologia Documental, fundamentam a Arquivologia nas pesquisas sobre práticas de gestão documental e tratamento de massas documentais acumuladas em arquivos (RODRIGUES, 2008).

A união da espécie documental com a função para a qual o documento foi elaborado resulta no tipo documental, que é objeto de estudo da Tipologia Documental, disciplina que ampliou a Diplomática para abranger o contexto, as atribuições, as competências, as funções e as atividades da entidade geradora/acumuladora dos documentos (BELLOTTO, 2002). Por meio da Tipologia Documental são analisados os vínculos de proveniência e organicidade do conjunto de documento, identificando primeiro os que participaram da mesma transação para, então, estabelecer as relações recíprocas dos procedimentos existentes, dos documentos entre si e dos documentos do mesmo grupo que executou outras transações (DURANTI, 1998).

Schellenberg enuncia que os documentos de arquivo “[...] se revelam significativos em relação à atividade e só secundariamente quanto ao assunto.” (SCHELLENBERG, 1980⁴ *apud* RODRIGUES, 2008). E os documentos seriam então “[...] as provas primordiais para as suposições ou conclusões relativas a essas atividades e às situações que elas contribuíram para criar, eliminar, manter

ou modificar.” (DURANTI, 1998, p. 2), desse modo, tanto as ações como os efeitos delas decorrentes.

A descrição é a função arquivística que busca identificar e explicar o contexto e o conteúdo dos documentos de arquivo para promover o acesso a eles, o que é feito por meio de representações precisas e adequadas de acordo com modelos predeterminados (CONSELHO INTERNACIONAL DE ARQUIVOS, 2000). Os documentos eletrônicos permitem o registro da representação neles mesmos com as linguagens de marcação como o XML, o que é baseado em uma estruturação prévia de acordo com tipologias documentais específicas (ORTEGA, 2008). O meio digital viabiliza a pré-representação, em que as etapas para apresentar, representar e indicar os pontos de acesso são realizadas no mesmo momento da construção do documento eletrônico.

A identificação consiste no “[...] ato de determinar a identidade do documento de arquivo, de caracterizar os caracteres próprios e exclusivos que conferem essa identidade.” (RODRIGUES, 2008, p. 65). Como função arquivística, a identificação é utilizada para delimitar a **identidade do documento de arquivo** por meio da atribuição de caracteres que lhe são próprios e exclusivos, em reunir as informações sobre o documento no contexto de sua produção com o uso da descrição dos elementos que constituem a sua identidade (RODRIGUES, 2011).

A vontade de manifestar e dar origem ao ato é que inicia o fluxo de criação do documento que, devido à sua relevância, demanda intervenção arquivística desde esse momento inicial (GAGNON-ARGUIN, 1999). A função arquivística de criação de informação consiste na atividade profissional para estabelecer as condições que assegurem qualidade, validade, credibilidade e sustentabilidade do sistema de informações, o que envolve a seleção da forma mais apropriada e dos relacionamentos entre os diferentes tipos de informação.

No âmbito da Diplomática Digital, iremos enfatizar o Projeto InterPARES (*International Research on Permanent Authentic Records in Electronic Systems*), uma iniciativa acadêmica com pesquisadores de diversos países, coordenada pela Universidade de British Columbia, no Canadá, que

busca desenvolver conhecimento teórico-metodológico para a preservação de longo prazo de documentos arquivísticos digitais autênticos.

O projeto é composto pela fase 1, desenvolvida de 1999 a 2001, que se concentrou na preservação da autenticidade de registros eletrônicos principalmente de documentos textuais produzidos e mantidos em bancos de dados e sistemas de gerenciamento de documentos (DURANTI; THIBODEAU, 2006). A fase 2, que avançou de 2002 a 2007, buscou desenvolver e articular os conceitos, princípios, critérios e métodos que possam assegurar a criação e manutenção de registros precisos e confiáveis e a preservação a longo prazo de registros autênticos no contexto das atividades artísticas, científicas e governamentais que são conduzidas. Já a fase 3, de 2007 a 2012, consistiu na aplicação da teoria e dos métodos de preservação digital desenvolvidos pelo InterPARES e outros esforços de pesquisa até hoje em planos de ação concretos para corpos de registros existentes que devem ser mantidos a longo prazo por arquivos.

Desde 2012 está em andamento o InterPARES *Trust*, que busca criar um arcabouço teórico e metodológico que apoie o desenvolvimento de redes integradas e consistentes de políticas, procedimentos, regulamentos, padrões e legislações que possam ser aplicados aos registros digitais criados, armazenados e acessados usando a Internet, de modo a aumentar a prover confiança pública e boa governança em benefício da economia digital e da memória persistente (JANSEN; DURANTI, 2013).

O InterPARES 1 definiu as características do registro eletrônico com base tanto na teoria arquivística como na teoria diplomática de forma complementar (DURANTI; THIBODEAU, 2006). Do ponto de vista arquivístico, os registros são partes de agregações, incluindo o exame das suas relações com outros registros, com as pessoas envolvidas em sua criação e com as atividades no curso das quais elas são criadas e usadas. Já a teoria diplomática propõe-se a examinar registros como itens, o que viabiliza a identificação das características incorporadas nos próprios registros.

InterPARES 1 também estabeleceu as seguintes características necessárias de um registro eletrônico: forma fixa, conteúdo completo e

inalterado; conteúdo imutável; ligações com outros registros por meio de código de classificação; contexto administrativo identificável; um autor, um destinatário e um redator; uma ação na qual o registro participa ou em que eles auxilia proceduralmente ou como parte do processo de decisão (DURANTI; THIBODEAU, 2006).

Em seguida, o InterPARES 1 adotou a hipótese a partir da Diplomática de que os registros tradicionais e os eletrônicos são semelhantes o suficiente para possibilitar a concepção de uma forma documental ideal, típica, com todos os elementos possíveis de um registro (DURANTI; THIBODEAU, 2006). A partir daí, foi criado um modelo (*template*) de registro eletrônico ideal com quatro seções: (1) Forma documental (elementos intrínsecos, nomes das pessoas que concorrem para a criação do registro, data cronológica, local de origem do registro, indicação e descrição da ação ou assunto, a atestação, e a declaração de validação; elementos extrínsecos: recursos de apresentação, assinaturas eletrônicas e selos, carimbo digital de data, sinais especiais); (2) Anotações (adições feitas ao registro após sua criação como parte de sua transmissão; adições feitas ao registro no curso de sua manipulação na qual o registro participa; adições feitas no registro no curso de sua gestão como um registro); (3) Contexto (categorização do registro em jurídico-administrativo, contexto de proveniência, contexto procedimental, contexto documentário e contexto tecnológico); e (4) Meio, que não está ligado indissociavelmente à mensagem por não armazenar o registro como tal.

4 Análise dos resultados

Glushko e McGrath (2005) apontam a surpreendente estabilidade do conceito de documento, o que seria explicado pela sua constituição e nítida utilidade: documento é um pacote independente de informações relacionadas por meio do qual as pessoas e as empresas interagem umas com as outras. A Engenharia de Documentos tem como objeto os documentos devido ao fato deles estruturarem a interface do aplicativo para o usuário, ou seja, eles são as interfaces para os serviços e os processos de negócios. A interface do documento é o que vai manter limpas e estáveis as relações entre os parceiros de negócio. Nesse

sentido, a utilidade do documento ultrapassa a comunicação de informações para alcançar o próprio consenso sobre as atividades de diferentes agentes e a execução dessas mesmas ações.

Os documentos transacionais podem ser assim classificados como documentos dispositivos na Diplomática, em que a ação e a documentação são concomitantes, ambas intrinsecamente ligadas, tendo o documento como o princípio e o fim da ação jurídica (PAOLI, 2010⁵ *apud* TOGNOLI, 2013).

Na Arquivologia, o conceito de Tipo Documental é o que apresenta características que o aproximam do documento transacional, em especial pelo vínculo estreito entre o documento e o procedimento que lhe dá origem, bem como pela homogeneidade de formato, conteúdo informativo e suporte (RODRIGUES, 2008). As relações entre espécie e tipo dos documentos são consideradas “[...] formas relativamente estáveis de enunciados em sua estrutura composicional, estilo e conteúdo temático.” (SANTOS; SOUSA; BARRANCOS, 2017).

O InterPARES 1 concluiu, a partir dos estudos de casos realizados pela pesquisa, que o meio não é um fator relevante na avaliação da autenticidade de um registro, de modo que a preservação de um registro eletrônico está ligada à capacidade em reproduzi-lo (DURANTI; THIBODEAU, 2006). Há então a diferença significativa entre a forma como um documento eletrônico é manifestado a uma pessoa e a forma em que ele é armazenado digitalmente, o que evidencia a diferença do documento digital de um tradicional, além de exigir uma descrição exata da natureza de um documento para que se possa verificar se ele tem sua identidade preservada quando sofre alterações em seu meio de armazenamento ou no formato de processamento (DURANTI; THIBODEAU, 2006).

E que a relação entre um registro eletrônico e um arquivo de computador pode ser de um-para-um, um-para-muitos, muitos-para-um ou muitos-para-muitos, na medida em que os termos *recorde file* nunca devem ser usados como intercambiáveis (DURANTI; THIBODEAU, 2006).

Foi evidenciado nos estudos de caso a importância de um conceito-chave: os atributos dos registros, ou seja, as características definidoras de cada

registro ou de um elemento de registro, sendo que este já uma parte constituinte da forma documental do registro (DURANTI; THIBODEAU, 2006). Os atributos explicitados fora do registro, a exemplo dos metadados, são essenciais para identificar objetos digitais que não possuem um conteúdo estável ou uma forma fixa.

Para demonstrar os conceitos das diferentes áreas é que elaboramos o quadro abaixo:

Quadro 4 - Comparativo entre conceitos

Engenharia de Documentos	Diplomática Digital e Arquivologia
<p>Documento transacional (ou centrado em dados): de uso primordial por aplicações de negócio, cuja descrição estrutural sozinha captura os aspectos mais importantes do conteúdo do documento (GLUSHKO; McGRATH, 2005).</p>	<p>Documento de arquivo: é produzido de forma involuntária no decurso de uma atividade, constitui-se na própria ação autodocumentada (MENNE-HARITZ, 1998). Documento é a informação registrada, em que informação é a mensagem destinada à comunicação no espaço ou no tempo (DURANTI; THIBODEAU, 2006).</p>
<p><i>Document components</i> (Componentes de documentos): são um conjunto de <i>building blocks</i> utilizados para montar semanticamente estruturas mais ricas e modelos de documentos (GLUSHKO; McGRATH, 2005).</p>	<p>Elemento de descrição: “[...] cada uma das categorias de informação que compõem a descrição normalizada de documentos.” (CONSELHO NACIONAL DE ARQUIVOS, 2006, p. 15).</p>
<p><i>Form convention</i>: é um entendimento compartilhado entre um autor independente e um leitor sobre como uma função específica pode ser comunicada usando uma forma específica (GEURTS, 2010). <i>Document model</i>: os documentos são implementações de modelos, levando em conta a necessidade da Engenharia de Documentos em definir modelos de diferentes tipos de documentos de uma forma rigorosa e inequívoca para que possamos automatizar o seu</p>	<p>Espécie documental: o modelo prescrito pelas normas jurídicas, com redação que segue um determinado padrão semântico e formatado de modo a ter validade e credibilidade (BELLOTTO, 2002). Tipo Documental: é a informação elaborada formalmente acompanhada de ação, que assim “[...] adquire a condição de mensagem, com intenção específica e assimilação possível.” (BARRETO, 1994, p. 8) Elementos internos do documento:</p>

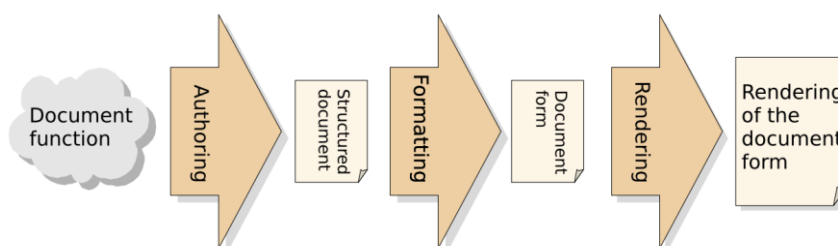
<p>processo de intercâmbio dentro ou entre aplicações (GLUSHKO; McGRATH, 2005).</p> <p><i>Schemadocument</i> (modelo lógico): abstrai as instâncias de documentos específicos e possibilita que o autor especifique restrições na estrutura de uma classe de documentos estruturados (GEURTS, 2010). Um documento sem um <i>Schema</i> é um mero saco de etiquetas (GLUSHKO; McGRATH, 2005).</p>	<p>também chamados de caracteres intrínsecos, substantivos ou de substância, referem-se ao conteúdo substantivo, o assunto propriamente dito, a natureza da sua proveniência e função. (BELLOTTO, 2002)</p> <p>Tanto os registros tradicionais como os eletrônicos são semelhantes o suficiente para possibilitar a concepção de uma forma documental ideal, típica, com todos os elementos possíveis de um registro. A partir daí, foi criado um modelo (<i>template</i>) de registro eletrônico ideal com quatro seções.</p> <p><i>Composition data</i>: dizem ao sistema quais <i>form data</i> e <i>content data</i> pertencem a qual documento (DURANTI; THIBODEAU, 2006).</p>
<p><i>Stylesheet</i> (folha de estilo): especifica a transformação de um documento estruturado para um formato de documento perceptível. (GEURTS, 2010).</p>	<p>Elementos externos do documento: também chamados de caracteres extrínsecos, físicos, de estrutura ou formais, referem-se à estrutura física e com a sua forma de apresentação (BELLOTTO, 2002).</p> <p><i>Form data</i>: dados que habilitam o sistema a reproduzir o documento na forma correta (DURANTI; THIBODEAU, 2006).</p>
<p>O documento transacional não é um objeto isolado: ele é ao lado do processo de negócio as duas faces da mesma realidade que é o desenvolvimento das atividades de uma organização (GLUSHKO; McGRATH, 2005).</p>	<p>Princípio da organicidade: os arquivos são o espelho da estrutura, funções e atividades da entidade que o produziu e acumulou (BELLOTTO, 2002).</p>
<p>Documentos como parte do processo de negócio, que é uma sequência de tarefas. (GLUSHKO, McGRATH, 2005)</p>	<p>Princípio do respeito à ordem natural, de modo que a ordem procedimental dos documentos reflita o procedimento administrativo que o origina. (HERRERA, 2003)</p>

Fonte: Adaptado de Glushko e McGrath (2005).

A partir da comparação podemos concluir que há muitas semelhanças e complementaridade no tratamento de documentos pela Engenharia de Documentos, Arquivologia e Diplomática Digital. Os conceitos utilizados não são exatamente idênticos, pois eles têm origem em áreas com objetivos diferentes, mas encontram-se muito próximos. Iremos adiante destacar dois pontos em que as abordagens são mais complementares entre si.

Os conceitos de espécie e tipo documental são centrais na Arquivologia e Diplomática Digital, mas não são utilizados conceitos semelhantes na Engenharia de Documentos. O paradigma da Engenharia de Documentos vai trabalhar a transformação da função em forma, o que é feito por meio de mudanças sequenciais em que o resultado da etapa anterior é usado como insumo da próxima fase (GEURTS, 2010). É o que se designa como cadeia de transformação, na qual as aplicações modernas utilizam o processo de três estágios (passos de transformação) de produção do documento:

Figura 1 – Cadeia de transformação composta por três passos de transformação.



Fonte: Geurts (2010, p. 29).

Na Autoria o conteúdo do documento é abstraído do seu formato para gerar um conjunto de dados explicitamente estruturado (*Structured document*), só que não é feita menção a modelos de referência utilizados, à semelhança do que seriam espécies ou tipos documentais. Na mesma linha, a Formatação é determinada pela aplicação do formato de documento adequado para um contexto de entrega específico, porém mais uma vez nos parece faltar conceitos como o de espécie e tipo documental para balizar essa aplicação. Afinal, para selecionar a estruturação dos dados e a forma do documento mais adequados é essencial que sejam confrontados com paradigmas existentes já consagrados pelo uso.

A espécie e o tipo documental carregam características próprias como uma estrutura de informações pensada para o exercício de determinada função por parte de uma organização, o que permite a identificação das séries documentais utilizadas. Com isso, a partir do modelo lógico e formato do documento identificado (espécie e tipo documental), é possível deduzir a função exercida. Ou, no sentido inverso: uma vez identificada a função, será possível deduzir qual a espécie e o tipo documental adequados no caso.

O ambiente digital disponibiliza uma variedade imensa de alternativas para organizar e representar os processos de negócios e suas transações, ou seja, poderíamos presumir que a abordagem prospectiva da Engenharia de Documentos abre caminho para criar formatos e estruturas lógicas inovadoras das informações a serem produzidas.

Mas, como é possível demonstrar, o caminho trilhado é exatamente o inverso. Primeiro, porque a abordagem da Engenharia de Documentos está toda centrada na figura do documento, este artefato com milênios de existência. Em segundo lugar, os conceitos articulados pela Engenharia de Documentos convergem para estudos da Diplomática que remontam há quase quatrocentos anos, e de Tipologia Documental que se desenvolveram nos últimos sessenta anos.

A utilização dos conceitos de documento, espécie e tipo documental como pontos de partida permitem que, uma vez combinados com a tecnologia existente, tenham seu potencial de representar informações multiplicadas exponencialmente. Portanto, os conceitos de espécie e tipo documentais trazem muitas possibilidades de contribuição significativa no desenvolvimento de pesquisas no âmbito da Engenharia de Documentos.

Outro ponto marcante identificado foi a proximidade dos **elementos de descrição** utilizados na Arquivologia e Diplomática Digital com os **componentes de documento** propostos pela Engenharia de Documentos. Um dos objetivos da Arquivologia é o acesso à informação pelo usuário, o que envolve instrumentos para organizar e representar as informações em níveis conceituais, tais como os elementos de descrição da Norma Brasileira de Descrição Arquivística (NOBRADE) e linguagens documentárias.

Já na Engenharia de Documentos a questão conceitual é uma etapa intermediária para o seu fim último, o qual consiste em executar aplicações de documentos, o que requer a definição dos seus componentes. O modelo básico de um documento é subdividido em dois tipos de componentes:

- a) componentes de conteúdo, que contém os valores de informação discreta;
- b) componentes estruturais, que são agregações dos anteriores, subdividindo-se em semânticas e de apresentação (GLUSHKO; McGRATH, 2005).

O componente estrutural de apresentação permite identificar a localização das informações no documento, o que muitas vezes pode indicar qual o tipo de significado daquela informação, ou seja, o componente de conteúdo do documento pode ser implicado pelo componente de apresentação desse mesmo documento.

Antona Copoulos, Karatzas, Krawczyk e Wiszniewski (2004) desenvolveram pesquisa em documentos textuais de relatos de prisioneiros da Segunda Guerra Mundial nos campos de concentração em que relacionaram os itens dos modelos de *layout* utilizados (componentes de apresentação) com os tipos de conteúdo (componentes de conteúdo) em cada item do *layout*. Com isso, puderam propor um Modelo de desenvolvimento de Ciclo de Vida de Documento Digital que atribui ao profissional especialista a tarefa de aceitação das informações somente nas hipóteses em que a classificação automatizada não apresentou resultados satisfatórios.

Assim, os componentes de documentos devem ser especificados e implementados não como definições de tipos de documentos somente, mas como artefatos participantes de uma arquitetura orientada a objetos para execução em aplicações (BRÜGGEMANN-KLEIN; WOOD, 2000).

Os componentes de documentos podem materializar nos sistemas eletrônicos das organizações as regras derivadas dos requisitos funcionais de um processo de negócio de modo a prover confiabilidade ao documento arquivístico ao estabelecer os formatos adequados para cada tipo de transação, estabelecer os elementos fixos de cada formato de documento e delimitar as possibilidades de criação do documento alinhadas ao processo padronizado (KETELAAR, 1997). Temos, assim, uma similaridade grande entre os componentes de documentos e

os elementos de descrição quanto ao tratamento conceitual dos dados e da informação, sendo que a maior diferença nos parece estar voltada à exigência de que os componentes atendam a determinados requisitos para compor aplicações.

No modelo contemporâneo de rede de relacionamentos entre organizações, os processos de negócio não ficam restritos a uma única organização, eles perpassam diversas delas em relações *business to business* (GLUSHKO; McGRATH, 2005). Para viabilizar as integrações, aqui consideradas como o controle compartilhado de dados ou processos de negócio entre aplicativos conectados ou fontes de dados, uma das opções é promover o *tightcoupling*, próprio para transações com altas taxas de dados. A sua principal desvantagem é a constante possibilidade de mudanças incontroláveis ou inesperadas nas interfaces (tais como o layout da tela, as estruturas de registro ou tabelas), situação que não é recomendada para integrações entre organizações.

Por isso, a Engenharia de Documentos e o InterPARES adotam a arquitetura fracamente acoplada (*loosecoupling*) para a integração entre aplicações de diferentes organizações por utilizar descrições mais abstratas de componentes dos documentos (DURANTI; THIBODEAU, 2006; GLUSHKO; McGRATH, 2005). Assim, justifica-se que a Engenharia de Documentos tenha uma abordagem centrada nos documentos pelo seu alto grau de estabilidade, o que contribui para a formação de consensos mais estáveis entre as organizações quanto aos significados dos documentos e os propósitos dos processos de negócios.

É muito significativo observar que, não obstante o volume imenso de dados e informações produzidos em documentos eletrônicos no mundo, as integrações entre os sistemas de organizações sejam baseadas de modo predominante nas descrições mais abstratas de componentes de documentos, que estruturam as regras dos processos de negócio envolvidos. Tais descrições representam um volume muito pequeno de dados e informações que são produzidos nessas organizações, mas a Engenharia de Documentos demonstra como as descrições abstratas desempenham papel crucial para que os processos de negócio possam se desenvolver com mais eficiência.

O segundo ponto é que a Diplomática Digital tem como um de seus principais objetivos a preocupação com a autenticidade e confiabilidade do documento eletrônico, o que é viabilizado por meio da utilização do modelo (*template*), o registro eletrônico ideal. InterPARES 2 definiu uma taxonomia de documentos em: documentos estáticos (não possibilitam alteração de conteúdo ou forma, mas apenas abertura, fechamento e navegação); documentos interativos (possuem conteúdo e/ou forma variáveis), podendo ser não-dinâmicos (com regras de conteúdo e forma que não variam) ou dinâmicos (DURANTI; THIBODEAU, 2006). Em seguida, analisou as implicações dos tipos de documentos conforme a sua capacidade em prover registros autênticos.

Na Engenharia de Documentos não foi identificada essa preocupação na criação de modelos de documentos nato-digitais para utilização em aplicações no futuro, pois seu objetivo maior foi de transformar documentos como interfaces de aplicações para implementar processos de negócios (GLUSHKO; McGRATH, 2005).

Nessa medida, a abordagem multidisciplinar nos parece muito útil para o preenchimento desta lacuna, em que a Diplomática Digital pode colaborar com a Engenharia de Documentos na aplicação de requisitos de autenticidade nos seus conceitos de *Document components*, *Form convention*, *Document model*, *Schema document* e *Stylesheet*.

5 Conclusões

No presente artigo pudemos abordar a evolução da Engenharia de Documentos, disciplina que ainda é pouco difundida no Brasil. O artigo de Quint, Nanard e André (1990), apresentado há 29 anos, dialoga sobre um conjunto significativo de desafios para a Engenharia de Documentos:

- a) as dificuldades de semântica devido aos diferentes usos do documento;
- b) a falta de elicitação de requisitos pelo autor do documento;
- c) a duplicidade de gramáticas que regem documentos (estrutura lógica e linguagem natural);
- d) os limites imprecisos entre os elementos estruturais (estrutura lógica, estrutura física e conteúdo);
- e) estrutura gráfica complexa para atender a leitura humana;

- f) início da escrita pelo autor em texto desestruturado;
- g) falta de conhecimento pelo editor do documento de sua estrutura lógica abstrata.

Os resultados da pesquisa revelaram como muitos desses desafios estão sendo superados e como as abordagens utilizadas se aproximam da Arquivologia e da Ciência da Informação. Isso aponta caminhos para a elaboração de estudos futuros sobre temas como documentos transacionais e documentos de arquivo, *schema documents* e elementos internos, *stylesheet* e elementos externos, procedimentos de análise e concepção do documento a partir de processos de negócio ou conjuntos documentais.

Retomamos a lição de Bachimont (2007) de que a Engenharia adentra uma dimensão original do conhecimento que inclusive excede a postura científica, por concretizar ações que a própria ciência não entende ainda. Stéphane Crozat (2012) resgatou o famoso artigo de Vannevar Bush de 1945, *As we may think*, elaborado no final da Segunda Guerra Mundial como reflexão sobre os meios técnicos para melhorar o armazenamento e recuperar informações dos grandes volumes documentais de produção científica da época. Passados 67 anos, Crozat (2012) inicia seu capítulo de livro com o título *As we may write* para mostrar que o digital também se tornou um outro meio para produzir o documento, em que impõe à escrita suas propriedades técnicas que transforma a própria natureza fundamental do documento.

No ambiente digital, a forma como o documento é escrito pode ter os seus efeitos potencializados ao ser produzido em sistemas eletrônicos, sejam eles para fins de recuperação da informação, seja para garantir a conformidade de execução dos processos de negócio. As contribuições da Engenharia de Documentos revelam mais uma necessidade de continuar a partir de conceitos utilizados na Arquivologia, Diplomática Digital e Ciência da Informação a partir das potencialidades e limitações da tecnologia, do que provocar propriamente uma ruptura nesse desenvolvimento.

Referências

ANTONACOPOULOS, Apostolos; KARATZAS, Dimosthenis; KRAWCZYK, Henryk; WISZNIEWSKI, Bogdan. The lifecycle of a digital historical document: structure and content. *In: ACM symposium on Document engineering, 2004, Milwaukee. Proceedings [...]*. Milwaukee: ACM, 2004. p. 147-154.

BACHIMONT, Bruno. **Ingénierie des connaissances et des contenus**: le numérique entre ontologies et documents. Paris: HermèsScience Publications, 2007.

BARRETO, Aldo de Albuquerque. A questão da informação. **São Paulo em Perspectiva**, São Paulo, v. 8, n. 4, 1994, p. 3-8.

BECKER, Christoph; RAUBER, Andreas; HEYDEGGER, Volker; SCHNASSE, Jan; THALLER, Manfred. A generic XML language for characterising objects to support digital preservation. *In: ACM symposium on Applied computing, 2008, Fortaleza. Proceedings [...]*. Fortaleza: ACM, 2008. p. 402-406.

BELLOTTO, Heloisa Liberalli. **Como fazer análise diplomática e análise tipológica em arquivística**: reconhecendo e utilizando o documento de arquivo. São Paulo: Arquivo do Estado: Imprensa Oficial do Estado, 2002. (Projeto Como Fazer).

BRÜGGEMANN-KLEIN, Anne; WOOD, Derick. **Document Engineering with Extensible Abstract Document Structures**. [s.l.]: HKUST Theoretical Computer Science Center Research Report, 2000.

BURKARD, Benjamin; VOGELER, Georg; GRUNER, Stefan. Informatics for historians: tools for medieval document XML markup, and their impact on the History-Sciences. **Journal of Universal Computer Science**, [s.l.] v. 14, n. 2, p. 193-210, 2008.

CONSELHO INTERNACIONAL DE ARQUIVOS. **ISAD(G)**: Norma geral internacional de descrição arquivística. 2. ed. Rio de Janeiro: Arquivo Nacional, 2000.

CONSELHO NACIONAL DE ARQUIVOS. **NOBRADE**: Norma Brasileira de Descrição Arquivística. Rio de Janeiro: Arquivo Nacional, 2006.

CROZAT, Stéphane. Chaînes éditoriales et rééditorialisation de contenus numériques. *In: CALDERAN, Lisette; LAURENT, Pascale; LOWINGER, Hélène; MILLET, Jacques (org.) .Le document numérique à l'heure du web*. Paris: ADBS, 2012. p. 179-220.

CROZAT, Stéphane. *Théorie de l'ingénierie des documents numériques*. Paris: UTC Formation, 2016. Disponível em: <<https://stph.scenari-community.org/doc/the.pdf>>. Acesso em: 08 nov. 2018.

DURANTI, Luciana. **Diplomatics: new uses for an old science**. Lanham: Scarecrow Press, 1998.

DURANTI, Luciana; THIBODEAU, Kenneth. The concept of record in interactive, experiential and dynamic environments: the view of InterPARES. **Archival Science**, Amsterdam, v. 6, n. 1, p. 13-68, 2006.

GAGNON-ARGUIN, Louise. La création. *In*: COUTURE, Carol (org.). **Les fonctions de l'archivistique contemporaine**. Sainte-Foy: Presses de l'Université du Québec, 1999. p. 69-101.

GEURTS, Jozef Petrus Theodorus Maria. **A document engineering model and processing framework for multimedia documents**. Eindhoven: SIKS Dissertation Series, 2010.

GLUSHKO, Robert J.; MCGRATH, Tim. **Document Engineering: analyzing and designing documents for business informatics and web services**. Massachusetts: MIT Press, 2005.

HERRERA, Antonia Heredia. *El principio de procedencia: y los otros principios de la archivística*. São Paulo: Associação de Arquivistas de São Paulo, 2003.

JANSEN, Adam; DURANTI, Luciana. The InterPARES Trust Project—Trust and Digital Records in an Increasingly Networked Society. *In*: GILLILAND, Anne; McKEMMISH, Sue; STANCIC, Hrvoje; SELJAN, Sanja; LASIC-LASIC, Jadranka (org.). **INFuture2013: Information Governance**. Zagreb: University of Zagreb, 2013. p. 63-68.

KETELAAR, Eric. Can we trust information?. **The International Information & Library Review**, London, v. 29, n. 3-4, p. 333-338, 1997.

LINS, Rafael Dueire. Meeting New Challenges in Document Engineering. **Journal of Universal Computer Science**, [s.l.], v. 17, n. 1, 2011, p. 1-2.

MENNE-HARITZ, Angelika. What can be achieved with archives? *In*: The concept of record: report from the Second Stockholm Conference on Archival Science and the Concept of Record, 1996, Stockholm. **Proceedings** [...]. Stockholm: Riksarkivet, 1998. p.11-24.

MUNSON, Ethan V. Abstract. *In*: ACM Symposium on Document engineering, 2001, Atlanta. **Proceedings** [...]. Atlanta: ACM, 2001.

ORTEGA, Cristina Dotta. Fundamentos da organização da informação frente à produção de documentos. **Transinformação**, Campinas, v. 20, n. 1, p. 7-15, 2008.

PIEZ, Wendell. Beyond the "descriptive vs. procedural" distinction. *In: Extreme Markup Languages*, 2001, Montreal. **Proceedings** [...]. Montreal: [s.n.], 2001. p. 141-172.

QUINT, Vincent; NANARD, Marc; ANDRÉ, Jacques. Towards document engineering. *In: EP 90 International Conference on Electronic Publishing, Document Manipulation & Typography*, 1990, Gaithersburg. **Proceedings** [...]. Gaithersburg: Cambridge University Press, 1990. p. 17-29.

RODRIGUES, Ana Célia. **Diplomática contemporânea como fundamento metodológico da identificação de tipologia documental em arquivos**. 2008. Tese (Doutorado em História Social) - Departamento de História, Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2008.

RODRIGUES, Ana Célia. Identificação: uma nova função arquivística? **Revista EDICIC**, [s.l.], v. 1, n. 4, p. 109-129, 2011.

SABLATNIG, Robert; HASSAN, Tamir. Abstract. *In: ACM Symposium on Document Engineering*, 2016, Vienna. **Proceedings** [...]. Vienna: ACM, 2016. p. 1-2.

SANTOS, Eliete Correia dos; SOUSA, Vancarde Brito; BARRANCOS, Jacqueline Echeverría. A interdisciplinaridade entre arquivologia e linguagem: a ciência à luz da perspectiva dialógica do discurso. **Biblionline**, João Pessoa, v. 13, n. 1, p. 71-81, 2017.

TOGNOLI, Natália Bolfarini. **A construção teórica da Diplomática**: em busca de uma sistematização de seus marcos teóricos como subsídio aos estudos arquivísticos. Tese (Doutorado em Ciência da Informação) – Programa de Pós-Graduação em Ciência da Informação, Universidade Estadual Paulista, Marília, 2013.

WHITE, Bebo. **Web Document Engineering**. [s.l.]: [s.n.], 1996. Disponível em: https://inis.iaea.org/collection/NCLCollectionStore/_Public/29/006/29006687.pdf. Acesso em: 12 nov. 2018.

WITSCHURKE, Karola. A Workbench for Document Processing. **International Journal of Information Theories & Applications**, Bulgária, v. 12, n. 2, p. 135-140, 2005.

Document Engineering and its relationships with Archivology, Digital Diplomatics and Information Science

Abstract: The Engineering of Documents is a Computer Science discipline that investigates the electronic documents. The present article consists of comparing and articulating the diverse information by the Engineering of Documents, Archivology, Digital Diplomatics and Information Science about documents and information. The methodology used is the documentary research of the literature that includes Digital Diplomatics, Documentary Typology and the archival characteristics of Creation and Description. The results were findings similarity between the document concepts used by the different areas of knowledge and the document conception from business processes or documentary sets. The main differences were: the lack of equivalence between the concepts of type and use of Archivology; the lack of treatment of issues of authenticity of anatomical documents of the Digital Diplomatics.

Keywords: Document Engineering. Archivology. Information Science. Digital Diplomatics. Documentary Typology.

Recebido: 12/02/2019

Aceito: 17/07/2019

¹ PÉDAUQUE, Roger T. Document: form, signandmedium, as reformulated for electronicdocuments. 2003. Terceira versão. *Apud* Crozat (2012).

² AUYANG, Sunny Y. Engineering – an endless frontier. Cambridge: Harvard University Press, 2006. *Apud* Bachimont (2007).

³ CARUCCI, Paola. Il documentcontemporaneo: diplomatica e criteridiedizione. Roma: La Nuova Italia Scientifica, 1987. *Apud* Tognoli (2013).

⁴ SCHELLENBERG, R. R. Documentos públicos e privados: arranjo e descrição. Rio de Janeiro: Editora FGV, 1980. *Apud* Rodrigues (2008).

⁵ PAOLI, Cesare. Diplomatica. Nuova edizione aggiornata da G.C. Bascapè con 220 disegni e facsimili. Firenze: Casa editrice Le lettere, Terza ristampa, 2010. *Apud* Tognoli (2013).