

# Documento e Contexto: uso de ontologias para representação semântica da produção documental arquivística.

Marieta Marks Löw<sup>1</sup>, Rafael Port da Rocha<sup>2</sup>, Mara Abel<sup>3</sup>, Luan Fonseca Garcia<sup>4</sup>

<sup>1</sup>Programa de Pós-Graduação em Ciência da Informação  
Universidade Federal do Rio Grande do Sul

<sup>2</sup>Faculdade de Biblioteconomia e Comunicação  
Universidade Federal do Rio Grande do Sul

<sup>3</sup>Instituto de Informática  
Universidade Federal do Rio Grande do Sul

<sup>4</sup>Instituto de Informática  
Universidade Federal do Rio Grande do Sul

marietalow@gmail.com, rafael.rocha@ufrgs.br, mara.abel@inf.ufrgs.br,  
luan.garcia@inf.ufrgs.br

**Abstract.** *The article describes a study on the representation of archival document production with BFO ontologies. Starting from a conceptual analysis in the Archival Science literature, the study analyzes the representation of the record and its production context using an oil exploration and production sharing contract. The concepts of record, producer, context and functions and the Principle of Provenance are analyzed from the perspective of Archival Science. As a result, an ontology is developed from IAO and BFO class specializations.*

**Resumo.** *O artigo descreve estudo sobre representação da produção documental arquivística com ontologias de BFO. Partindo de análise conceitual na literatura arquivística o estudo analisa a representação do documento arquivístico e seu contexto de produção utilizando um contrato de partilha de exploração e produção de petróleo. São analisados os conceitos de documento arquivístico, produtor, contexto e funções e o Princípio da Proveniência a partir do olhar da Arquivologia. Como resultado, uma ontologia é desenvolvida a partir de especializações de classes de IAO e BFO.*

## 1. Introdução

Compreender a produção documental pelo olhar da proveniência arquivística significa compreender a entidade produtora, e com isso suas funções e atividades do negócio que deram origem aos documentos produzidos, ou seja, compreender o documento em seu contexto de produção. Desse modo, representar os documentos e seus relacionamentos orgânicos pode significar uma melhora na capacidade de organização do conhecimento desse domínio, bem como o desenvolvimento de ferramentas mais efetivas para gestão e recuperação da informação.

A produção documental de uma instituição compõe um conjunto de documentos que permitem, de maneira única, compreender a atuação social da entidade e as mudanças ocorridas ao longo do tempo. A gestão dessa documentação é feita, tradicionalmente, seguindo princípios basilares da Arquivologia, entre eles o Princípio de Proveniência, que

estabelece a manutenção do conjunto de documentos gerados em uma mesma entidade produtora de forma a que fiquem reunidos na instituição arquivística com a organização que lhe foi dada originalmente [Schellenberg 2006]. Desse modo, olhados em conjunto, os documentos permitem compreender funções e atividades da instituição produtora [Cook, 2012], tornando-se importante fonte informacional do fazer da instituição.

O escopo deste trabalho refere-se à análise do documento arquivístico pela perspectiva da sua produção, considerando o contexto e a proveniência arquivística, de modo a desenvolver soluções para uma melhor compreensão e representação do documento em sistemas de informações. A proveniência arquivística estabelece a vinculação dos documentos ao seu produtor e ao seu contexto de produção. Trata-se de uma pesquisa exploratória de cunho qualitativo. Investiga de que modo os conceitos da Arquivologia podem ser utilizados na representação ontológica da produção documental, utilizando a ontologia Information Artifact Ontology (IAO) e a ontologia de topo Basic Formal Ontology (BFO). Analisa conceitualmente, através de pesquisa bibliográfica e documental, os conceitos e princípios arquivísticos e de IAO e BFO para a representação da produção documental arquivística. Abrange documentos produzidos no âmbito do contrato de exploração e produção de petróleo da área de Libra como exemplo para avaliar as possibilidades de representação ontológica dos resultados obtidos na análise conceitual.

Ao usar ontologias para representação da produção documental considerando os princípios arquivísticos espera-se que haja uma melhoria da capacidade semântica da representação ao serem explicitados os relacionamentos entre os documentos e o produtor, e entre os documentos e as atividades desenvolvidas. Buscou-se então uma ontologia que já tivesse entidades voltadas à representação de documentos para subsidiar o trabalho. IAO é uma ontologia de entidades de informação, derivada do trabalho da Ontology for Biomedical Investigations (OBI). A partir de entidades de BFO foram desenvolvidas classes para representação de objetos digitais e entidades informacionais.

A escolha do contrato de Libra como exemplo de estudo para a produção documental justifica-se por ser uma empresa de grande porte e com processos complexos de produção documental. O setor de produção de petróleo tem uma legislação bastante rigorosa e específica quanto à comprovação das atividades às agências de regulação como a ANP (Agência Nacional do Petróleo, Gás Natural e Biocombustíveis) e o IBAMA (Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis). Desenvolver e propor ferramentas de análise e classificação de documentos a partir das funções do negócio que os geraram, considerando as regulamentações que constituem esse negócio, proporciona ganhos em agilidade e eficiência no acesso à informação.

Este estudo traz a representação semântica do documento pela perspectiva da proveniência arquivística, que decorre de princípio basilar da Arquivologia, chamado de Princípio da Proveniência. A proveniência arquivística é base para normas de descrição de arquivos, como ISAD(G) e ISDF, e para construção de sistemas de gestão arquivística de documentos, como ISO 15489. A proposição de uma ontologia que estende BFO e IAO para a representação da produção documental a partir dos princípios arquivísticos possibilita compreender o alcance e a compatibilidade de BFO e IAO em representar o documento arquivístico. Além disso, traz uma representação ontológica para documentos que está alinhada com os princípios e as estratégias da Arquivologia para descrição e gestão de documentos. BFO foi adotada neste projeto por ser indicada como padrão para o

desenvolvimento de ontologias na indústria pela associação Industrial Ontology Foundry<sup>1</sup> oferecendo módulos de ontologias previamente desenvolvidas que foram integradas àquela desenvolvida neste projeto. Um desses módulos corresponde à IAO que define um arcabouço conceitual maduro para a representação de entidades informacionais, alinhado com o objetivo desta pesquisa. Em especial, a IAO descreve documentos como entidades ontologicamente dependentes que representam informação “sobre” alguma entidade do domínio. Representar a relação de dependência ontológica da entidade de informação e ser referente a uma entidade material se aproxima da conceitualização arquivística, como descrita neste artigo na seção 2, e tornam a IAO uma estrutura conceitual adequada para a representação da relação da produção documental com as entidades das ontologias de domínio baseadas na BFO desenvolvidas no projeto ou anteriores a ele.

Outros trabalhos abordam documento arquivístico e ontologia, embora não tendo como foco de discussão os conceitos de documento arquivístico e proveniência arquivística, tais como estudos sobre a representação de esquemas de classificação para arquivos multimídia [Vsesviatska *et al* 2021], sobre a representação ontológica para fundos documentais [Pandolfo *et al* 2018] e sobre representação de papéis ligados à registro de autoridade arquivística [Daquino *et al* 2014]. O Conselho Internacional de Arquivos (ICA) vem desenvolvendo na última década um modelo conceitual e uma ontologia específicos para a representação de conjuntos documentais arquivísticos e seu contexto de produção, o Records in Context, RiC-CM (modelo conceitual) e RiC-O (ontologia) [ICA 2019, 2021]. Apesar de tratar a representação de documentos arquivísticos, a ontologia RiC-O ainda é uma minuta, e não está ancorada em nenhuma ontologia de topo [Llanes-Padrón 2017], como BFO, que é foco da discussão deste artigo. Uma ontologia derivada de IAO, a Document Acts Ontology [Brochhausen *et al* 2013; Smith 2014] se aproxima da questão da produção de documentos através do ato documental e dos poderes deônticos que um documento exerce sobre os indivíduos, do ponto de vista jurídico. Não é com foco nos conceitos arquivísticos, mas pode ser analisada complementarmente. PROV-O [W3C 2013] é uma ontologia para representar proveniência de informação gerada em diferentes sistemas e sobre diferentes contextos, podendo ser especializada para modelar proveniência em diferentes domínios e aplicações.

Nossa pesquisa iniciou com um levantamento documental utilizando princípios e técnicas da Arquivologia. O objetivo foi compreender a produção documental de um consórcio de partilha de produção e exploração de petróleo, como o Consórcio de Libra, sob a perspectiva arquivística. Foi, também, desenvolvido estudo da representação da produção documental a partir da análise conceitual de entidades derivadas de BFO e da teoria arquivística. Com os resultados obtidos a partir destes dois enfoques foi possível desenvolver a representação ontológica da produção documental de um consórcio de produção e exploração de petróleo com o uso de IAO e BFO. A seção 2 apresenta conceitos sobre documento e documento arquivístico na literatura das áreas de Ciência da Informação e da Arquivologia. Na seção 3 são trazidos os resultados da análise da literatura acerca dos conceitos arquivísticos de produção documental, produtor, contexto e proveniência arquivística. Na seção 4 são apresentados os resultados do levantamento documental e de contexto da produção documental de um consórcio de produção e exploração de petróleo. Na seção 5 são analisadas as entidades de IAO e BFO frente aos conceitos arquivísticos

---

<sup>1</sup> [www.industrialontologies.org](http://www.industrialontologies.org)

levantados nas seções 2 e 3 e debatidos os resultados da representação ontológica da produção documental de Libra apresentada na seção 4.

## 2. Documento e Documento Arquivístico

No âmbito da Ciência da Informação (CI), Paul Otlet e Suzanne Briet, importantes autores da escola documentalista, conceituaram documento reforçando seu carácter de registro e evidência. Para Otlet (2018), o documento é uma forma de representar coisas materiais, ou imateriais como imagens intelectuais e abstratas. Nesse sentido o documento é sobre alguma coisa. Ressalta-se nessa concepção a questão da materialidade do objeto, bem como a perspectiva do documento como evidência e testemunho, e como representação de algo, material ou imaterial. Briet inaugura uma progressiva superação da visão iluminista e universalista de Otlet. Os documentos passam a ser portadores de um “status de evidência”, podendo ser utilizados para consulta, estudo ou prova [Gugliotta 2017].

Essa concepção foi atualizada por Buckland (1991) ao nomeá-lo “informação-como-coisa”. Buckland parte do conceito de informação para chegar ao documento, como uma informação tangível pela sua materialização em algum suporte. Ao analisar a conhecida discussão de Briet sobre um antílope tornado documento ao ser trazido de seu habitat para um zoológico e ser estudado por cientistas, Buckland infere os elementos que podem determinar quando um objeto se torna um documento: materialidade (objetos físicos e sinais físicos); intencionalidade (o objeto é intencionalmente tratado como evidência); processamento (os objetos são transformados em documentos); e posição fenomenológica (o objeto é percebido como documento) [Buckland 1997, p. 806]. O autor considera que há aspectos físicos (suporte e técnicas de produção), cognitivos (sentido e contexto da produção) e sociais (contexto cultural e papéis desempenhados) a serem observados para compreensão do documento [Buckland 2018].

Na perspectiva da Arquivologia o termo documento, no sentido amplo, equivale ao termo utilizado por autores da CI, enquanto registro de informação, já o termo documento arquivístico é usado com sentido mais especializado. Não se trata apenas de uma adjetivação, mas de um conceito chave que é o objeto de estudo da área, com características próprias de produção e uso [Rondinelli 2013]. Na língua inglesa o uso do termo documento (*document*) para representar o documento de arquivo não é tão frequente na literatura quanto o termo documento arquivístico (*record* em inglês). Documentos (*records*) são:

Todos os livros, papéis, mapas, fotografias ou outras espécies documentárias, independentemente de sua apresentação física ou características, expedidos ou recebidos por qualquer entidade pública ou privada no exercício de seus encargos legais ou em função das suas atividades e preservados ou depositados para preservação por aquela entidade ou por seus legítimos sucessores como prova de suas funções, sua política, decisões, métodos, operações ou outras atividades, ou em virtude do valor informativo dos dados neles contidos. [Schellenberg 2006, p. 41]

Autores mais recentes abordam questões como suporte e contexto, temáticas que se tornaram relevantes com a introdução da tecnologia e uso de computadores na produção de documentos. Luciana Duranti (2002, p. 22) conceitua documento arquivístico (*record*) como: “[...] *any document created by a physical or juridical person in the course of practical activity as an instrument and a by-product of it*”. Já G. Yeo (2007, p. 337) traz a questão do documento como representação de algo que aconteceu: “[...] *records can be characterized*



*as persistent representations of activities, created by participants or observers of those activities or by their authorized proxies” (grifo do autor).*

A discussão principal é acerca da natureza dos arquivos. Há um entendimento de que a acumulação dos documentos arquivísticos é parte do fazer de um determinado produtor, sua acumulação é orgânica, ou seja, os documentos de arquivo são produzidos, recebidos e acumulados a partir de uma atividade, por um propósito. Isso diferencia os documentos arquivísticos, ou seja, o arquivo, enquanto conjunto documental, dos documentos em coleções, como em bibliotecas ou centros de documentação.

### 3. Produção e Produtor, Contexto e Proveniência Arquivística do Documento

A partir de análise dos conceitos de documento e documento arquivístico na literatura, alguns termos se destacam na tentativa de compreender o fenômeno da produção documental: há um **produtor**, pessoa física ou entidade coletiva, que produz, recebe e acumula os documentos; há um **contexto** de produção, normas, regulamentos, procedimentos que definem a produção dos documentos de um dado produtor e sua organização como um fundo documental; há uma ação, ou seja, **funções** e atividades ou transações realizada pelo produtor, que originam a produção dos documentos; e, por fim há o **documento arquivístico**, evidência material da ação realizada pelo produtor, preservada para fins de prova ou informação. Sendo o documento produzido em razão das atividades de um produtor, sua produção ocorre como resultado de um processo que acontece no tempo e no espaço, desse modo, cada arquivo tem características próprias que ficam claras quando os documentos são percebidos no seu conjunto. Destaca-se, também, o **Princípio da Proveniência**, como definido pela arquivologia, que estabelece a relação de proveniência arquivística dos documentos, ou seja, a vinculação dos documentos ao seu produtor, e a relação de organicidade, que vincula os documentos entre si, e com seu contexto de produção

Em língua inglesa o termo usado para **produtor** é *creator*. A norma de descrição de arquivos ISAD(G), em seu glossário de termos, conceitua produtor como a entidade coletiva, família ou pessoa que produziu, acumulou ou manteve documentos na gestão de sua atividade coletiva ou pessoal, termo que não deve ser confundido com colecionador. O mesmo glossário também faz distinção entre produtor e autor, já que o segundo é o responsável intelectual de um documento. [Conselho Internacional de Arquivos 2000].

Para o projeto InterPARES 3, **contexto** (*context*) é definido como a estrutura na qual o documento é criado, usado e mantido [InterPARES 3 Project 2021]. Na pesquisa, foram identificados alguns tipos de contextos possíveis na produção de documentos arquivísticos: o **contexto jurídico-administrativo**, o sistema legal e organizacional ao qual pertence o produtor de documentos arquivísticos; o **contexto administrativo**, as estruturas, funções e procedimentos do ambiente organizacional ao qual pertence o produtor de documentos; o **contexto documental**, o fundo ao qual o documento arquivístico pertence e a estrutura interna desse mesmo fundo; o **contexto de proveniência**, a entidade produtora, seu mandato ou missão, estrutura e funções; além do **contexto tecnológico**, as características de hardware, software e outros componentes de um sistema eletrônico computacional no qual os documentos arquivísticos são criados.

A produção dos documentos é consequência de ações que são realizadas pelo produtor de acordo com regras, normas e procedimentos próprios. Em uma organização, a produção documental é consequência de **funções** e atividades que fazem parte do seu processo de negócio. No âmbito da gestão arquivística de documentos, a norma ISO 15489

recomenda a elaboração de um sistema de classificação funcional de atividades de negócio, [International Organization for Standardization 2001], que é base para o desenvolvimento de instrumentos de gestão documental, como os esquemas de classificação. A análise com a finalidade de desenvolver essa classificação funcional identifica todas as atividades de uma organização e as localiza na estrutura de sua missão ou propósito. Na sua forma plenamente desenvolvida, essa classificação produz uma representação das funções, que são decompostas em subfunções, por sua vez decompostas em procedimentos operacionais e assim sucessivamente.

Segundo a ISDF, norma arquivística para descrição de funções, uma função corresponde a

qualquer objetivo de alto nível, responsabilidade ou tarefa prescrita como atribuição de uma entidade coletiva pela legislação, política ou mandato. Funções podem ser decompostas em conjuntos de operações coordenadas, tais como subfunções, procedimentos operacionais, atividades, tarefas ou transações [Conselho Internacional de Arquivos 2007, p. 10].

A mesma norma diz que a “compreensão das funções das entidades coletivas é essencial para uma percepção integral da proveniência arquivística. Os documentos arquivísticos têm uma relação fundamental com funções. Eles são o produto direto do exercício de funções” [Conselho Internacional de Arquivos 2007, p. 30]. Assim, o conjunto de todos os **documentos arquivísticos** de um produtor forma um fundo documental. Esses conjuntos são arranjados e descritos, de acordo com a norma de descrição arquivística ISAD-G, de forma hierárquica, em vários níveis, onde conjuntos são formados por subconjuntos e assim sucessivamente na hierarquia. A norma Nobrade, que aplica ISAD-G no Brasil, prevê como níveis para a descrição de um fundo documental: seção, série, dossiê ou processo e item documental. [Conselho Nacional de Arquivos (Brasil) 2006] Práticas atuais recomendam que essa organização hierárquica de arranjo tenha como orientação a visão funcional da organização. Tais práticas arquivísticas buscam compreender e representar o funcionamento e a produção documental do produtor.

A partir da análise conceitual é possível perceber algumas especificidades da representação da produção de documentos arquivísticos. A relação de **proveniência** é uma delas, e está relacionada a origem do relacionamento entre os documentos arquivísticos e seu produtor, e entre os documentos e seu contexto de produção. Estes relacionamentos são compreendidos através do que a teoria arquivística denominou de Princípio da Proveniência, definido como: “Princípio básico da arquivologia segundo o qual o arquivo produzido por uma entidade coletiva, pessoa ou família não deve ser misturado aos de outras entidades produtoras. Também chamado princípio do respeito aos fundos” [Arquivo Nacional 2005, p. 136].

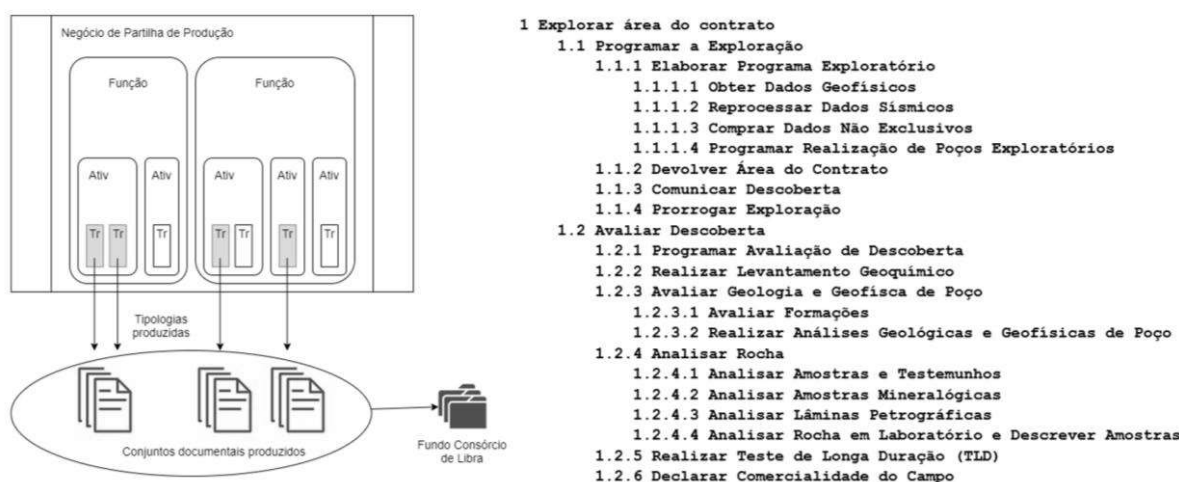
#### 4 Representação da Produção Documental Arquivística

Para a representação da produção documental foi feito um levantamento documental utilizando princípios e técnicas da Arquivologia, buscando compreender a questão sob a perspectiva arquivística, para então aplicar essa representação em ontologia por meio da extensão de IAO e BFO. O estudo foi feito no domínio da exploração e produção de óleo e gás, mais especificamente, em contexto de partilha da exploração da produção do petróleo. O estudo tomou como base consórcios de produção e exploração de petróleo, regidos pelo modelo de contrato de partilha e documentos regulatórios publicados pela ANP, tendo o

Consórcio de Libra como referência. O estudo ateve-se ao mapeamento dos níveis superiores e médio da hierarquia de funções, já que os níveis inferiores implicam em aspectos particulares da operacionalização de cada consórcio.

O levantamento documental foi realizado com base em informações divulgadas publicamente e compreendeu as diversas etapas de exploração e produção de petróleo. Teve como base publicações, notícias e sites sobre Libra e, principalmente, leis, modelo de contrato, resoluções e normas infralegais brasileiras, especialmente da agência brasileira de petróleo, que regulam vários aspectos da produção, exploração e partilha de petróleo. A partir desses documentos foi possível identificar e especificar as funções de negócio de alto e médio nível, assim como as tipologias documentais relacionadas, que são aquelas exigidas pela agência reguladora. Nesses documentos podemos encontrar os aspectos chaves sobre os contextos jurídico-administrativo, administrativo, documental e de proveniência. A metodologia para identificação das funções de negócio teve como base a norma ISO 15489 e o Manual DIRKS, principal referência para o desenvolvimento da norma ISO [Macintosh and Real 2007; Healy 2010]. Tais normativos são considerados referências técnicas para a área de arquivos, e indicados para a representação da estrutura funcional de entidades produtoras de documentos. São identificadas e representadas, em instrumentos próprios da área, as funções de mais alto nível da organização, seguidas daquelas de nível intermediário até as atividades e transações de nível operacional.

O estudo desenvolveu uma hierarquia com processos de negócio onde foram representadas funções, atividades e transações de negócio, com a identificação das respectivas tipologias produzidas em cada nível. No nível mais alto foram identificadas sete funções: explorar área do contrato; desenvolver produção; produzir petróleo e gás; gerir reservatórios; partilhar produção; apoiar exploração e produção (E&P); gerir serviços corporativos. As funções são decompostas em subfunções, procedimentos operacionais, atividades e transações (tarefas), contemplando os cinco níveis hierárquicos indicados na norma de descrição de funções ISDF.



**Figura 1 Relação entre as funções de negócio e a produção documental**

Com o levantamento de funções e atividades, foi possível verificar as tipologias produzidas e a relação com as atividades da entidade produtora. Além das sete funções de negócio identificadas no nível mais alto, o quadro de funções tem também 21 subfunções, 40 procedimentos operacionais, 35 atividades, e quatro tarefas. Foram relacionadas também 138 tipologias documentais e 31 documentos normativos. A partir da hierarquia funcional foi





uma folha de papel (*Material Information Bearer*) ocorre através de um carregador da informação (*Information Carrier*), como as marcas de tinta de uma impressora. Em IAO o carregador da informação é uma qualidade que permite a concretização do conteúdo no portador. Sendo uma qualidade, depende de uma entidade específica, o portador, e deixa de existir se a primeira também cessa sua existência.

Documento arquivístico, como visto acima, se caracteriza por ser produzido e acumulado no decurso de atividades de um dado produtor e mantido em conjunto com outros documentos arquivísticos para fins informacionais e probatórios. Desse modo, considerando estas especificidades, foi proposto, para o estudo, a especialização da entidade documento (*Document*), de IAO, em documento arquivístico (*record*), conforme Figura 2.

Até aqui o termo documento foi analisado como um continuante. Entretanto, a ação de documentar aponta para o fato de que o documento tem uma existência no tempo e no espaço, significando que ele surge a partir de um processo de produção pela entidade produtora. BFO contempla a representação do processo de produção do documento. A ligação entre os conceitos aparece pela entidade processo planejado (*Planned Process*), um termo reutilizado em IAO.

É possível perceber o documento arquivístico como um resultado em um processo planejado através do qual são realizadas determinadas funções de negócio, as quais tem como resultado registros materializados em documentos arquivísticos. Um processo planejado (*Planned Process*) é aquele que realiza um plano, o qual é uma concretização de uma especificação de plano (*Plan Specification*). Na ontologia desenvolvida (Figura 2), a classe *Business Planed Process* representa processos planejados de negócio que realizam a concretização de planos de gestão documental (*Recordkeeping Plan*). Esses processos são classificados em processos que compõem a hierarquia de funções de negócio da organização (*Business Function*, *Business Subfunction*, *Business Procedure* e *Business Task*), de acordo com a norma ISDF; e em processos de gestão documental (*Recordkeeping Process*), que regulam a criação e a gestão do documento arquivístico. *Record Creation Process* (Figura 2) é um processo de gestão documental que determina que um documento seja criado por pessoas devidamente autorizadas, com todos os elementos constituintes requeridos pelo sistema jurídico e como resultado de uma função de negócio da organização (*Business Planned Process*) ou suas subdivisões.

Já a representação do produtor é contemplada com BFO por meio da entidade Organização (*Organization*), utilizada em OBI, através de um papel (*Role*) de produtora de documentos. Em BFO, papéis (*Role*) são entidades realizáveis através de processos, as quais se concretizam em entidades materiais. Para OBI, uma organização (*Organization*) é uma entidade capaz de desempenhar papéis, a qual possui membros e tem um conjunto de regras organizacionais. Na ontologia (Figura 2), dois tipos de papéis foram especificados: *Records Creator Role*, que representa a organização no papel de produtora de documentos, e *Organization Role*, que representa pessoas na organização com a autoridade para gerar documentos. Por exemplo, Consórcio Libra é uma organização, que, como tal, possui o papel (*Records Creator Role*) de produtora de documentos, através do qual realiza o processo de criação de documento.

Na Arquivologia o conjunto dos documentos produzidos pelo produtor em função de suas atividades é chamado de fundo documental. A norma de descrição arquivística ISAD-G prevê como níveis para a descrição de um fundo documental: seção, série, dossiê ou processo e item documental. Na ontologia criada (figura 2), a classe *fundo (fond)* está

vinculada a classe *conjunto documental* (*record set*). Cada conjunto documental é parte de uma hierarquia de conjuntos representadas na ontologia como “*part of*”, indicando o nível superior de agregação dos documentos. Isso permite a representação do contexto documental.

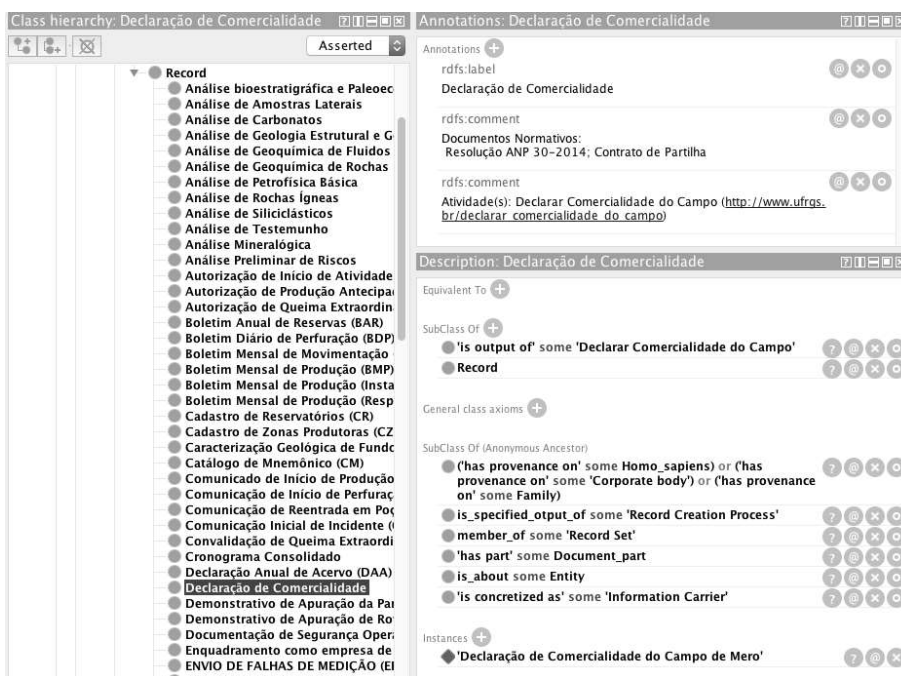


Figura 3: Representação da tipologia Declaração de Comercialidade

A figura 3 mostra a classe *Record* e suas subclasses. Essas subclasses representam as tipologias documentais, isto é, os tipos de documentos que são produzidos por subdivisões das funções de negócios (como *Business Procedure* e *Business Task*). No caso, representam as tipologias documentais produzidas pelas funções de negócio de um consórcio de partilha da exploração e produção de petróleo levantadas conforme apresentado na seção 4. A figura 3 exemplifica a tipologia Declaração de Comercialidade, que é um documento (*Document*) que é saída do (*is output of*) procedimento de negócio de Declarar Comercialidade de Campo e, também, do processo de gestão documental (*Record Creation Process*).



Figura 4: Representação das atividades geradoras de documentos

A funções de negócio de um consórcio de partilha da exploração e produção de petróleo levantadas conforme apresentado na seção 4, por sua vez, foram incluídas como subclasse de um processo planejado de negócio, agrupadas na hierarquia estabelecida pela metodologia arquivística. A Figura 4 apresenta subclasses identificadas para a classe tarefa de negócio, conforme a hierarquia funcional estabelecida pela metodologia arquivística. É possível relacionar a atividade “Analisar Geoquímica de Poço” como parte da atividade

“Analisar Rocha em Laboratório e Descrever Amostras”. A relação *member of* indica que o documento é membro de um conjunto documental.

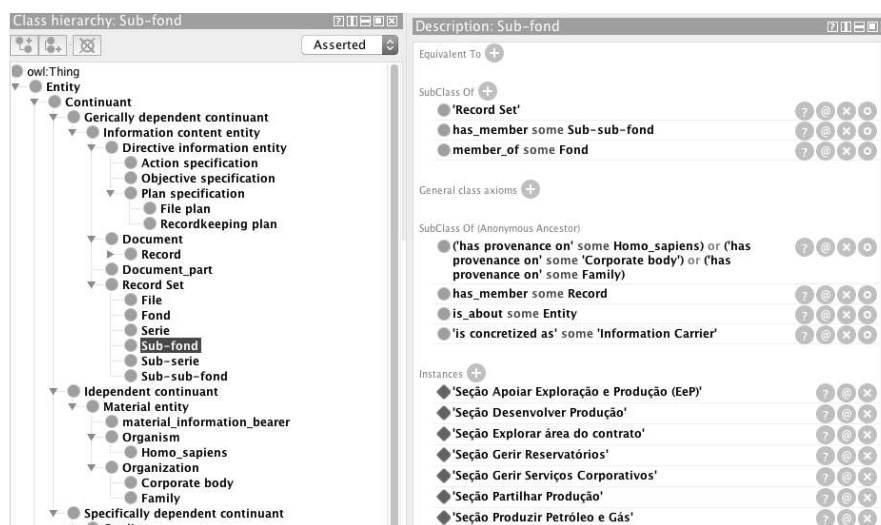


Figura 5: Representação dos conjuntos documentais

Os documentos gerados durante os processos de negócio são parte de conjuntos documentais (Figura 5). Na ferramenta os conjuntos documentais foram representados como instâncias das classes de conjuntos documentais. Assim, o Consórcio de Libra gera uma documentação representada pela instância “Fundo Consórcio de Libra”, na classe fundo (*fond*), que está vinculada a classe conjunto documental (*record set*). Cada conjunto documental é parte de uma hierarquia de conjuntos representadas na ferramenta como “*part of*”, indicando o nível superior de agregação dos documentos. Isso permite a representação do contexto documental. Na Figura 5 vemos as instâncias da classe Seção (*Sub-fond*) e a relação da Seção com outros conjuntos documentais, como ser parte de um Fundo (*Fond*), e ter como partes Subseções (*Sub-sub-fond*).

## 6. Considerações

Este trabalho apresentou resultados de análise conceitual no domínio da produção documental arquivística e sua aplicação para a representação ontológica de conceitos arquivísticos. Com a análise fundamentada na bibliografia arquivística e aplicação prática com levantamento da produção documental na área de óleo e gás, o estudo verificou como representar ontologicamente o documento arquivístico e seu contexto de produção, a partir de princípios da Arquivologia.

Com a realização do estudo foi possível abordar a representação da produção documental arquivística a partir de uma realidade específica, o contrato de partilha da produção de petróleo, para verificação das análises teóricas. Com isso pois possível constatar a aplicabilidade de BFO e IAO para representação desse domínio específico que é a produção do documento arquivístico. As entidades de BFO, como ontologia de topo permitem a representação da produção documental, do produtor e do contexto de maneira ampla. O uso de ontologias derivadas de BFO, como IAO, permitem a representação do documento em aspectos mais específicos. Para características próprias da teoria arquivística foi possível utilizar as entidades sugeridas a partir da análise teórica.

Na ontologia desenvolvida por meio da especialização de classes de BFO, o documento arquivístico é interpretado pela perspectiva de que ocorre como resultado de um

processo planejado (*Planned Process*), que é realizado pela organização (*Organization*) no papel (*Role*) de produtora de documentos, na concretização do que foi especificado em planos (*Plan Specification*). Esse processo planejado é parte do negócio da organização, cuja hierarquia funcional, formada por funções, subfunções, procedimentos operacionais e atividades (conforme a norma ISDF) é representada em BFO por processos maiores que possuem processos como partes.

Nessa ontologia, regras de negócio e de gestão documental são registradas em especificações de planos (*Plan Specification*) e que regulam a execução de funções organizacionais (*Planned Process*). As funções de negócio são realizadas por pessoas, que, quando atribuídas a papéis (*Roles*), fazem com que essas pessoas sejam institucionalmente reconhecidas na autoria dos documentos.

Com a especificação de classes e propriedades implementadas em OWL, alguns aspectos levantados na discussão teórica foram verificados. A solução proposta foi validada com a representação do conjunto documental proveniente do contrato de Libra, onde foi possível explicitar o produtor, os documentos produzidos, as ações que geraram os documentos e a relação de proveniência entre eles. Por exemplo, a importância do relacionamento entre os conjuntos documentais, onde um é parte integrante de outro conjunto, este um conjunto maior composto de vários conjuntos. Como o caso do Fundo, que agrega todos os demais. É esta explicitação de relação que permite a representação do contexto documental. Na ferramenta foi possível descrever esse relacionamento com a propriedade “*part of*”. Essa mesma relação também possibilita a explicitação do contexto de proveniência, com a ligação entre uma atividade e um procedimento operacional, por exemplo.

O contrato de partilha de petróleo serviu como referência para a representação da produção documental a partir de proposta de extensão de IAO e BFO para abarcar a produção de documentos arquivísticos. Entretanto, sendo a atividade petrolífera uma área cujos resultados dependem sobremaneira do uso de dados e informações, a geração de conhecimento teórico e de soluções tecnológicas para gestão e recuperação de informação para o setor petrolífero ganha importância com a possibilidade de aplicação concreta dos desenvolvimentos da pesquisa na indústria. A atividade de exploração e produção de petróleo é uma atividade complexa, e com normativos legais e corporativos muito específicos. O mapeamento dos processos que originam a produção documental, e os usos e gestão dos documentos produzidos são recorrentes em corporações de grande porte como as da indústria petrolífera. O uso de ferramentas semânticas para a representação dos documentos e do contexto de produção pode facilitar e ampliar o acesso a informações estratégicas contidas no acervo documental na indústria.

**Agradecimentos:** Este projeto foi financiado pela CAPES (Código 001) e pela PETROBRAS, através do termo de cooperação do Projeto Geologia Digital: Busca integrada de dados geocientíficos heterogêneos. Agradecemos a equipe do projeto pela colaboração no desenvolvimento e dados para o estudo.

## Referências

- Almeida, Mauricio Barcellos. (2020) Ontologia em Ciência da Informação: teoria e método. CRV Curitiba.
- Arquivo Nacional. (2005) Dicionário brasileiro de terminologia arquivística: Publicações Técnicas. Arquivo Nacional, Rio de Janeiro.



- Briet, Suzanne. (1970) O que é documentação. Universidade Federal Fluminense, Instituto de Artes e Comunicação Social, Niterói.
- Brochhausen, Mathias; Almeida, Mauricio; Slaughter, Laura. (2013) Towards a formal representation of document acts and resulting legal entities. In: Verlag (org.). Johanssonian Investigations. p. 120–139.
- Buckland, Michael K. (1991) Information as thing. In: *Journal of the American Society for Information Science*, v. 42, n. 5, pages 351–360.
- Buckland, Michael K. (1997) What is a ‘document’? In: *Journal of the American Society for Information Science*, v. 48, n. 9, pages 804–809.
- Buckland, Michael K. (1998) What is a digital document? In: *Document numérique*, v. 2, n. 2, pages 221–230.
- Buckland, Michael K. (2014) Documentality beyond documents. In: *The Monist*, v. 97, n. 2, pages 179–186.
- Buckland, Michael K. (2020) Document Theory. In: *Knowledge Organization*, v. 45, n. 5, pages 425–436.
- Conselho Internacional de Arquivos. (2000) ISAD(G): Norma geral internacional de descrição arquivística. 2a Ed. Arquivo Nacional, Rio de Janeiro.
- Conselho Internacional de Arquivos. (2007) ISDF: Norma internacional para descrição de funções. Conselho Internacional de Arquivos, Paris.
- Conselho Internacional de Arquivos (2019) Records in Context: a conceptual model for archival description (Consultation Draft v0.2). [https://www.ica.org/sites/default/files/ric-cm-0.2\\_preview.pdf](https://www.ica.org/sites/default/files/ric-cm-0.2_preview.pdf).
- Conselho Internacional de Arquivos (2019) Records in Contexts Ontology (ICA RiC-O) version 0.2. [https://www.ica.org/standards/RiC/RiC-O\\_v0-2.html](https://www.ica.org/standards/RiC/RiC-O_v0-2.html).
- Conselho Nacional de Arquivos (Brasil). (2006) NOBRADE: Norma Brasileira de Descrição Arquivística. Arquivo Nacional, Rio de Janeiro.
- Cook, Terry. (2012) A ciência arquivística e o pós-modernismo: novas formulações para conceitos antigos. In: *InCID: Revista de Ciência da Informação e Documentação*, v. 3, n. 2, p. 3–27, Ribeirão Preto.
- Daquino, Marilena; Peroni, Silvio; Tomasi, Francesca; Vitali, Fabio. (2014) Political roles ontology (PROles): Enhancing archival authority records through semantic web technologies. *Procedia Computer Science*, v. 38, p. 60–67, 2014.
- Duranti, L.; Eastwood, T.; Macneil, H. (2002) Preservation of the Integrity of Electronic Records. Springer Netherlands, Dordrecht.
- Gugliotta, Alexandre Carlos. (2017) Pensando e repensando o documento In: *Revista Ibero-Americana de Ciência da Informação*, v. 10, n. 2, pages 314–331.
- Healy, Susan. (2010) ISO 15489 Records Management: Its development and significance. In: *Records Management Journal*, v. 20, n. 1, pages 96–103.
- Heredia Herrera, Antonia. (1991) Archivística general: teoría y práctica. 5a. Ed. Servicio de Publicaciones de la Diputación de Sevilla, Sevilla.

- International Organization for Standardization – ISO (2001) ISO/TR 15489-2:2001(E). Geneva.
- InterPARES 3 Project. International Terminology Database, [http://www.interpares.org/ip3/ip3\\_terminology\\_db.cfm](http://www.interpares.org/ip3/ip3_terminology_db.cfm).
- Llanes-Padrón, Dunia; Pastor-Sánchez, Juan Antonio. (2017) Records in contexts: the road of archives to semantic interoperability. *Program*, v. 51, n. 4, p. 387–405,
- Macintosh, Stephen; Real, Lynne. (2007) DIRKS: Putting ISO 15489 to Work. In: *Information Management Journal*, v. 41, n. 2, pages 50–56.
- Otlet, Paul. (2018) Tratado de documentação: o livro sobre o livro, teoria e prática. Briquet de Lemos, Brasília.
- Pandolfo, Laura; Pulina, Luca; Zieliński, Marek. (2018) ARKIVO: An ontology for describing archival resources. In: Felli, Paolo; Montali, Marco. (Org.) *CEUR Workshop Proceedings: 33rd Italian Conference on Computational Logic*. CEUR-WS, Bolzano. p. 112–116.
- Rondinelli, Rosely Curi. (2013) O documento arquivístico ante a realidade digital: Uma revisão conceitual necessária. FGV, Rio de Janeiro.
- Schellenberg, T. R. (2006) Arquivos Modernos, princípios e técnicas. 6a Ed. FGV, Rio de Janeiro.
- Smith, Barry. (2014) “Document Acts”, Institutions, Emotions, and Group Agents: Contributions to Social Ontology. Konzelmann-Ziv, A.; Schmid, H. B. Springer, Dordrecht. pages 19–31.
- Smith, Barry and Ceusters, Werner. (2015) Aboutness: Towards foundations for the information artifact ontology. In: 2015, CEUR Workshop Proceedings. : CEUR-WS.
- Smith, Barry *et al.* Basic Formal Ontology 2.0 - Specification and User’s Guide. At: <https://github.com/BFO-ontology/BFO/raw/master/docs/bfo2-reference/BFO2-Reference.pdf>.
- Tognoli, Natalia; Guimarães, José Augusto Chaves. (2019) Provenance as a knowledge organization principle. In: *Knowledge Organization*, v. 46, n. 7, pages 558–568.
- Vsesviatska, Oleksandra; Tietz, Tabea; Hoppe, Fabian; Sprau, Mirjam; Meyer, Nils; Dessì, Danilo; Sack, Harald. (2021) ArDO: an ontology to describe the dynamics of multimedia archival records. In *Proceedings of the 36th Annual ACM Symposium on Applied Computing (SAC '21)*. Association for Computing Machinery, New York, NY, USA, 1855–1863. DOI:<https://doi.org/10.1145/3412841.3442057>
- W3C (2013) PROV-O: The PROV Ontology. W3C Recommendation 30 April 2013. <https://www.w3.org/TR/prov-o/>.
- Yeo, Geoffrey. (2007) Concepts of Record (1): Evidence, Information, and Persistent Representations In: *The American Archivist*, v. 70, n. 2, pages 315–343.