

## **A PRODUÇÃO CIENTÍFICA SOBRE PRESERVAÇÃO DIGITAL: LEVANTAMENTO BIBLIOGRÁFICO E DOCUMENTAL EM BASES DE DADOS E REPOSITÓRIOS INSTITUCIONAIS**

**Resumo:** Frente a crescente necessidade de criação e sistematização de práticas relacionadas à preservação digital este trabalho identifica, descreve e analisa um conjunto de registros obtidos sobre preservação digital em bases de dados e repositórios institucionais com abrangência nacional e internacional, sendo elas: *Science Direct*, *Scopus*, *Web of Science* e LISA, além da Biblioteca Digital Brasileira de Teses e Dissertações do IBICT, BDTD da USP, Biblioteca Digital da UNICAMP, BDTD da UNESP, SciELO e Brapci. Objetiva identificar os períodos e as fontes em que se possa obter o maior número de registros acerca do tema preservação digital no intuito de promover um arranjo quantitativo sobre o assunto, que subsidie pesquisas e ações futuras. Configura-se como uma pesquisa exploratória, de natureza bibliográfica e documental. Constitui o corpus de análise registros obtidos no período entre 2003 a 2013. A análise dos registros pautou-se no método de levantamento bibliográfico e documental. Verifica que as bases Brapci e *Web of Science* apresentaram respectivamente, no panorama nacional e internacional, o maior número de registros. No cenário internacional, o período de 2009 a 2011 se destacou na produção de estudos, enquanto que as revistas científicas se sobressaíram como principais fontes de informação. Conclui-se, através do pequeno volume de resultados constatado em certas bases e repositórios, que se torna indispensável o desenvolvimento de mais iniciativas nacionais neste domínio.

**Palavras-chave:** Preservação digital. Levantamento bibliográfico e documental. Base de dados. Repositório institucional. Ciência da Informação.

**Danilo Formenton**  
Mestre em Ciência, Tecnologia e Sociedade pela  
Universidade Federal de São Carlos - UFSCar.  
formenton.danilo@gmail.com

**Luciana de Souza Gracioso**  
Doutora em Ciências da Informação pela  
Universidade Federal Fluminense, Mestre em  
Ciência da Informação PUC Campinas.  
lgracioso@yahoo.com.br

### **SCIENTIFIC PRODUCTION ABOUT DIGITAL PRESERVATION: BIBLIOGRAPHIC DOCUMENTARY SURVEY IN DATABASES AND INSTITUTIONAL REPOSITORIES**

**Abstract:** Regarding the growing need about creation and systematization of digital preservation practices, this paper aims to identify, describe and analyze a set of records about digital preservation obtained in national and international databases and institutional repositories as: ScienceDirect, Scopus, Web of Science, LISA, Brapci and Scielo, in addition to Brazilian Digital Library of Theses and Dissertations of IBICT, USP, UNICAMP and UNESP. The analyses aimed to identify the sources and the periods in which is possible get the highest number of registrations about digital preservation topic in order to promote a quantitative arrangement on this subject, to subsidize bibliographic researchs and future actions. This work is classified as a bibliographic and documentary exploratory research. The analysis sampling is constituted of records obtained between 2003 and 2013. This Analysis was based on the bibliographic and documentary method. It was verified that Brapci and Web of Science databases presented respectively the largest number of records in national and international scope. On the international scenario, the period 2009-2011 was highlighted in the production of studies, and scientific journals were the major source of information. In conclusion, because of the small volume of results recovered in certain databases and repositories, it becomes essential developing more national initiatives.

**Keywords:** Digital Preservation. Bibliographic and documentary research. Databases. Institutional repository. Information Science.

## 1 INTRODUÇÃO

A preservação digital vem nos últimos anos sendo alvo de investigação pela comunidade científica nacional e internacional. Os estudos publicados em periódicos nacionais no campo da Ciência da Informação, geralmente pautam-se numa discussão mais conceitual sobre o tema, relatando tal processo como sendo o caminho mais adequado atualmente para garantia da preservação e acesso utilizável, por longo prazo, ao conteúdo das informações digitais. Russel e Sergeant (1999), por exemplo, consideram a preservação digital como sendo o armazenamento, a manutenção e o acesso dos objetos digitais por longo prazo, usualmente como consequência da aplicação de uma ou mais estratégias de preservação digital, incluindo a preservação de tecnologia, a emulação tecnológica ou a migração de dados.

Entretanto, dos trabalhos publicados no contexto do campo da Ciência da Informação, são poucos aqueles que se dedicam exclusivamente ao desenvolvimento de um levantamento bibliográfico e documental em bases de dados e repositórios institucionais focado na análise dos resultados, a fim de identificar os períodos e tipos de fontes para auxiliar na obtenção de dados e informações em futuras pesquisas sobre tal temática. Além disso, baseado em estudos recentes produzidos neste campo, podemos afirmar ainda que comparado a alguns países europeus como, por exemplo, Portugal e outros que são referência no desenvolvimento de pesquisas e projetos nesta temática, ainda é singela a contribuição brasileira no que diz respeito tanto a própria realização de estudos científicos sobre preservação digital como a criação de iniciativas em prol da preservação e do acesso utilizável de recursos de informação em meio digital por longo período.

A partir deste cenário, por meio da análise dos resultados recuperados no levantamento realizado em determinadas bases e repositórios de abrangência nacional e internacional em estudos científicos, este trabalho possui o propósito de identificar os períodos e os tipos de fontes onde se possa obter o maior número de estudos e informações acerca dos diversos aspectos que envolvem a preservação digital, a fim de colaborar nas futuras pesquisas e projetos nacionais no contexto da área da Ciência da Informação. Sendo assim, a presente investigação trata-se de uma pesquisa exploratória, onde a coleta de dados baseou-se em pesquisa bibliográfica e documental e, apresentou como característica geral de configuração, o levantamento bibliográfico e documental em bases de dados e repositórios institucionais.

## 2 PROCEDIMENTOS METODOLÓGICOS

Através da análise qualitativa de conteúdo de resumos e da consulta das respectivas palavras-chave presentes em artigos publicados no período entre 2002 a 2012 em periódicos científicos nacionais e portugueses, foram elaboradas e elencadas algumas expressões de busca em língua portuguesa e inglesa com ou sem operadores de busca booleanos de intersecção e união. Os motivos para a inclusão de artigos portugueses na investigação dos dados para o trabalho tem seu fundamento em função deste país possuir, por exemplo, reconhecimento nacional e internacional na literatura como por órgãos governamentais pelas diversas iniciativas e projetos de preservação, a promoção de eventos, um conjunto significativo de trabalhos publicados em periódicos científicos do campo da Ciência da Informação além da perspectiva futura de talvez ocorrer ações conjuntas de cooperação quanto ao tema.

A partir de um grupo de 36 expressões de busca, realizou-se um levantamento bibliográfico e documental entre o período de 2003 a 2013 sobre preservação digital em bases de dados e repositórios institucionais com abrangência nacional e internacional em estudos científicos, a fim de obter um panorama mais atual dos registros sobre esta temática de investigação. Os registros após serem recuperados, primeiramente foram organizados e inseridos em um conjunto de quadros que, em seguida, foram analisados e retirados dados considerados mais pertinentes para a elaboração das diversas tabelas expostas neste trabalho. Para o tratamento e análise dos resultados com o intuito de elaborar as tabelas, utilizou-se os recursos disponíveis no *software Microsoft Office Excel* versão 2010, sendo que para o arredondamento de um número para um número especificado de dígitos (um, dois ou três, por exemplo) aplicou-se a função *ARRED*.

Em princípio, para apresentar um panorama internacional sobre o assunto, as bases ScienceDirect (Elsevier), Scopus (Elsevier), Web of Science (Thomson Reuters) e Library and Information Science Abstracts – LISA (ProQuest), foram selecionadas em razão da sua abrangência mundial, pelo grande volume de artigos e de periódicos em diversas áreas, por disponibilizar referências, resumos, textos integrais ou dados de citações além de inúmeros recursos para diferentes estratégias de busca e refinamento de resultados. Posteriormente, para também expor um cenário nacional quanto ao tema, ocorreu à escolha das bases e repositórios Biblioteca Digital de Teses e Dissertações (BDTD) da Universidade de São Paulo (USP), Biblioteca Digital (BD) da Universidade Estadual de Campinas (UNICAMP) e Biblioteca Digital de Teses e Dissertações (BDTD) da Universidade Estadual Paulista “Júlio de Mesquita Filho” (UNESP), devido a tais instituições de ensino superior estarem entre as melhores no país

e serem os centros de maior produção de conhecimento do estado de São Paulo. Estas bases e repositórios serviram também para uma melhor apuração dos resultados obtidos das instituições do estado de São Paulo na Biblioteca Digital Brasileira de Teses e Dissertações (BDBTD) do Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT). Além disto, a Base de Dados Referencial de Artigos de Periódicos em Ciência da Informação (Brapci) foi consultada por ser reconhecida como instrumento importante na divulgação, disponibilização e análise de estudos nacionais na área da Ciência da Informação e, também a Scientific Electronic Library Online (SciELO) por ser gratuita, apresentar uma grande quantidade de periódicos, possuir abrangência nacional e em diversos outros países.

Durante o processo, a escolha do emprego ou não simultâneo de expressões em língua portuguesa e inglesa com ou sem o uso de operadores booleanos assim como dos recursos e filtros de busca disponíveis nas bases e repositórios, fundamentou-se nas opções onde se obteve o maior número de resultados. Embora durante todo o processo tenha ocorrido à indicação do período a ser abrangido para o levantamento, identificaram-se em certos casos registros recuperados não inclusos neste período. O processo foi desenvolvido entre os meses de junho e julho do ano de 2013. Deste modo, foram recuperados e expostos nas tabelas apenas dados parciais quanto aos registros referentes ao ano de 2013. Ademais, o acesso aos *websites* oficiais de grande maioria das bases de dados ocorreu por meio do Portal de Periódicos da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes), com a inserção de seus respectivos nomes no campo Buscar base. Nas bibliotecas digitais e na Brapci, o acesso foi realizado diretamente por meio de seus *websites* oficiais.

### **3 RESULTADOS OBTIDOS NAS BASE DE DADOS COM ABRANGÊNCIA INTERNACIONAL EM ESTUDOS CIENTÍFICOS**

#### **3.1 Base de dados ScienceDirect<sup>1</sup>**

No processo de levantamento realizado nesta base de dados, obteve-se um conjunto elevado de resultados (882 registros), sendo que deste volume 75,2% correspondeu à categoria de assunto *All Sciences* (663 registros) e 24,8% a categoria *Social Sciences* (219 registros).

---

<sup>1</sup> Adotaram-se o recurso *Search* e posteriormente *All Sources*, sendo que todas as expressões de busca foram inseridas no campo *Search* na opção *Abstract, Title, Keywords* em *All Sources*. Para cada expressão, na opção *Subject*, empregou-se a categoria de assunto *All Sciences (Life Sciences, Health Sciences, Physical Sciences and Engineering e Social Sciences & Humanities)* numa primeira busca e, depois, somente a categoria *Social Sciences* na segunda busca. Em ambas as buscas realizadas, se utilizaram os itens *Journals* e *All Books* na opção de *Include*. Além disso, aplicou-se *All sources* na opção *Source* e o valor *2003 to Present* em *Date Range*.

Desta forma, a base de dados *ScienceDirect* se constituiu no presente processo como sendo a fonte onde se recuperou o quarto maior número de estudos produzidos no mundo sobre preservação digital. Assim, pode-se estabelecer na Tabela 1, a partir do número total de resultados obtidos nesta base para estas categorias, a distribuição ao longo de onze anos da quantidade e subsequente frequência da ocorrência dos registros relacionados aos assuntos de preservação digital.

**Tabela 1** - Distribuição da quantidade e frequência da ocorrência de registros obtidos para as categorias de assunto no período entre 2002-2013, base de dados *ScienceDirect*.

<b>Ano</b>	<b>Número total de registros na categoria de assunto <i>All Sciences (AS)</i> e frequência em porcentagem</b>	<b>Número total de registros na categoria de assunto <i>Social Sciences (SS)</i> e frequência em porcentagem</b>	<b>Número total de registros nas categorias de assuntos <i>All Sciences (AS)</i> e <i>Social Sciences (SS)</i> e frequência em porcentagem</b>
2002	3 (0,5%)	3 (1,4%)	6 (0,7%)
2003	41 (6,2%)	22 (9,9%)	63 (7,1%)
2004	52 (7,8%)	20 (9%)	72 (8,1%)
2005	34 (5,1%)	2 (0,9%)	36 (4,1%)
2006	53 (8%)	17 (7,7%)	70 (7,9%)
2007	71 (10,7%)	20 (9%)	91 (10,2%)
2008	77 (11,6%)	22 (9,9%)	99 (11,1%)
2009	67 (10,1%)	30 (13,5%)	97 (10,9%)
2010	49 (7,4%)	13 (5,9%)	62 (7%)
2011	70 (10,5%)	18 (8,1%)	88 (9,9%)
2012	79 (11,9%)	31 (14%)	110 (12,4%)
*2013	70 (10,5%)	24 (10,8%)	94 (10,6%)
<b>TOTAL</b>	666 (75% do número total de registros para ambas as categorias de assunto)	222 (25% do número total de registros para ambas as categorias de assunto)	888 (100%)

**Fonte:** Elaborado pelos autores.

Através da tabela foi possível observar que no número de registros obtidos somente para a categoria de assunto *Social Sciences*, o período entre os anos de 2011 a 2013 (73 registros, 32,9%) demonstrou um elevado índice de produtividade de trabalhos durante estes onze anos. Ademais, os anos de 2012, 2009 e 2013 (85 registros, 38,3%) se destacaram com as maiores quantidades e subsequentes frequências de ocorrência de registros. Apesar da existência de dados parciais referentes ao ano de 2013, torna-se notável o crescimento progressivo de iniciativas nos últimos anos, em especial a partir do ano de 2009, por parte do campo das Ciências Sociais, evidenciando assim que a preservação digital e os desafios que se

\* Foram recuperados e apresentados apenas dados parciais quanto aos registros referentes ao ano de 2013. Portanto, tal marcação se seguirá ao longo das demais tabelas, sendo que em todas as ocorrências de análises referentes a 2013 deverão ser levadas em conta este resultado parcial.

propõem a resolver vêm sendo investigados pelas áreas do conhecimento que compreendem este campo.

Além disto, a Tabela 2 presente abaixo expressa o número total de resultados obtidos para os tipos de informação, documento ou conteúdo denominados *Journal*, *Book* e *Reference Work* nas categorias de assunto *All Sciences* e *Social Sciences*, com indicação da quantidade e frequência da ocorrência de registros.

**Tabela 2** - Distribuição da quantidade e frequência da ocorrência de registros obtidos para os diversos tipos de informação nas categorias de assunto, base de dados *ScienceDirect*.

Número total de registros para <i>Journal</i> nas categorias de assunto <i>All Sciences</i> (AS) e <i>Social Sciences</i> (SS), além de frequência em porcentagem		Número total de registros para <i>Book</i> nas categorias de assunto <i>All Sciences</i> (AS) e <i>Social Sciences</i> (SS), além de frequência em porcentagem		Número total de registros para <i>Reference Work</i> nas categorias de assunto <i>All Sciences</i> (AS) e <i>Social Sciences</i> (SS), além de frequência em porcentagem	
AS	SS	AS	SS	AS	SS
586 (74,6%)	199 (25,4%)	90 (81,8%)	20 (18,2%)	16 (53,3%)	14 (46,7%)
785 (AS e SS, 100%)		110 (AS e SS, 100%)		30 (AS e SS, 100%)	

Fonte: Elaborado pelos autores.

Nesta tabela constatou-se que do número total de registros obtidos para *Journal* (785 registros), *Book* (110 registros) e *Reference Work* (30 registros), respectivamente 25,4% (199 registros), 18,2% (20 registros) e 46,7% (14 registros) corresponderam à categoria *Social Sciences*. Deste modo, as áreas que integram o campo das Ciências Sociais como, por exemplo, a área da Ciência da Informação, podem através deste levantamento assumir as revistas científicas, os livros e as obras de referência presentes na base *ScienceDirect*, nesta sequência, como grandes fontes de informação para o desenvolvimento de pesquisas e práticas no âmbito do processo de preservação digital.

### 3.2 Base de dados Scopus<sup>2</sup>

Com relação ao levantamento efetuado nesta base de dados, obteve-se um conjunto grande de resultados (2683 registros), sendo que deste volume 64,3% equivaleu à categoria de assunto *All Sciences* (1726 registros) e 35,7% a categoria *Social Sciences e Humanities* (957

<sup>2</sup> Adotaram-se o recurso *Search* e posteriormente *Document search*, sendo que todas as expressões de busca foram inseridas no campo *Search for* na opção *Article Title, Abstract, Keywords*. Para cada expressão na opção *Subject Areas*, selecionaram-se todas as áreas (*Life Sciences, Health Sciences, Physical Sciences e Social Sciences & Humanities*) numa primeira busca e, depois, somente a área *Social Sciences & Humanities* na segunda busca. Em ambas as buscas efetuadas, escolheu-se a opção *All* em *Document Type*. Além disto, aplicou-se o valor *Published 2003 to Present* em *Date Range* no campo de *Limit to*.

registros). Sendo assim, a base de dados *Scopus* se estabeleceu no presente processo de levantamento, como sendo a fonte onde se obteve o terceiro maior número de estudos produzidos no mundo sobre a temática investigada. À vista disto, pode-se estabelecer na Tabela 3, a partir do número total de resultados obtidos nesta base para essas categorias, a distribuição ao longo de dez anos da quantidade e subsequente frequência da ocorrência dos registros associados aos assuntos de preservação digital.

**Tabela 3** - Distribuição da quantidade e frequência da ocorrência de registros obtidos para as categorias de assunto no período entre 2003-2013, base de dados *Scopus*.

Ano	Número total de registros na categoria de assunto <i>All Sciences (AS)</i> e frequência em porcentagem	Número total de registros na categoria de assunto <i>Social Sciences &amp; Humanities (SS&amp;H)</i> e frequência em porcentagem	Número total de registros nas categorias de assunto <i>All Sciences (AS)</i> e <i>Social Sciences &amp; Humanities (SS&amp;H)</i> , além da frequência em porcentagem
2003	28 (1,6%)	16 (1,7%)	44 (1,6%)
2004	132 (7,6%)	42 (4,4%)	174 (6,5%)
2005	115 (6,7%)	51 (5,3%)	166 (6,2%)
2006	191 (11,1%)	117 (12,2%)	308 (11,5%)
2007	141 (8,2%)	99 (10,3%)	240 (8,9%)
2008	168 (9,7%)	87 (9,1%)	255 (9,5%)
2009	221 (12,8%)	124 (13%)	345 (12,9%)
2010	253 (14,7%)	131 (13,7%)	384 (14,3%)
2011	198 (11,5%)	115 (12%)	313 (11,7%)
2012	244 (14,1%)	158 (16,5%)	402 (15%)
*2013	35 (2%)	17 (1,8%)	52 (1,9%)
<b>TOTAL</b>	1726 (64,3% do número total de registros para ambas as categorias de assunto)	957 (35,7% do número total de registros para ambas as categorias de assunto)	2683 (100%)

**Fonte:** Elaborado pelos autores.

A partir da tabela verificou-se que no número total de registros obtidos apenas para a categoria de assunto *Social Sciences & Humanities*, o período entre os anos de 2009 a 2011 (370 registros, 38,7%) possuiu um alto índice de produtividade de trabalhos durante estes dez anos. Os anos de 2012, 2010 e 2009 (413 registros, 43,2%) se destacaram com as maiores quantidades e frequências de ocorrência de registros. Assim, os resultados advindos da base *Scopus* reforçam de certa forma o que foi constatado na base *ScienceDirect*, principalmente em relação a produção de trabalhos, identificada nos anos de 2009 e 2012.

Além disso, pode-se elaborar na Tabela 4 com base no número total de resultados obtidos nesta base para as categorias de assunto *All Sciences* e *Social Sciences & Humanities*, a distribuição da quantidade e frequência da ocorrência de registros relacionados aos assuntos de preservação digital nos diversos tipos de documentos.

**Tabela 4** - Distribuição da quantidade e frequência da ocorrência de registros obtidos para os diversos tipos de documento nas categorias de assunto, base de dados *Scopus*.

Número total de registros para os diversos tipos de documento na categoria de assunto <i>All Sciences (AS)</i> e frequência em porcentagem		Número total de registros para os diversos tipos de documento na categoria de assunto <i>Social Sciences &amp; Humanities (SS&amp;H)</i> e frequência em porcentagem		Número total de registros para os diversos tipos de documento nas categorias de assunto <i>All Sciences (AS)</i> e <i>Social Sciences &amp; Humanities (SS&amp;H)</i> , além de frequência em porcentagem	
<i>Article</i>	661 (38,3%)	<i>Article</i>	558 (58,31%)	<i>Article</i>	1219 (45,43%)
<i>Article in Press</i>	3 (0,17%)	<i>Article in Press</i>	1 (0,1%)	<i>Article in Press</i>	4 (0,15%)
<i>Conference Paper</i>	880 (50,98%)	<i>Conference Paper</i>	269 (28,11%)	<i>Conference Paper</i>	1149 (42,83%)
<i>Conference Review</i>	44 (2,55%)	<i>Conference Review</i>	19 (1,99%)	<i>Conference Review</i>	63 (2,35%)
<i>Editorial</i>	14 (0,81%)	<i>Editorial</i>	13 (1,36%)	<i>Editorial</i>	27 (1,01%)
<i>Note</i>	6 (0,35%)	<i>Note</i>	4 (0,42%)	<i>Note</i>	10 (0,37%)
<i>Review</i>	116 (6,72%)	<i>Review</i>	93 (9,72%)	<i>Review</i>	209 (7,79%)
<i>Short Survey</i>	2 (0,12%)	<i>Short Survey</i>	0 (0%)	<i>Short Survey</i>	2 (0,07%)
<b>TOTAL</b>	1726 (100%)	<b>TOTAL</b>	957 (100%)	<b>TOTAL</b>	2683 (100%)

**Fonte:** Elaborado pelos autores.

Na tabela pode ser observado que no número total de registros obtidos somente na categoria de assunto *Social Sciences & Humanities*, identificou-se altas quantidades e frequências de ocorrência de registros nos documentos *Article*, *Conference Paper* e *Review* (920 registros, 96,14%). Desta maneira, embora se tratando do campo das Ciências Sociais e Humanidades, novamente os artigos publicados em revistas científicas que abrangem as áreas que compõem estes campos, se sobressaem como principais fontes de informação na base *Scopus* para a obtenção de dados ou informações sobre preservação digital.

Pode-se também realizar na Tabela 5 a partir do número total de resultados obtidos nesta base para as categorias de assunto *All Sciences* e *Social Sciences & Humanities*, a distribuição da quantidade e frequência da ocorrência de registros relacionados aos assuntos de preservação digital no conjunto dos cinco primeiros países com maior volume de registros além de Brasil.

**Tabela 5** - Distribuição da quantidade e frequência da ocorrência de registros obtidos para os cinco primeiros países com maior volume de registros além de Brasil nas categorias de assunto, base de dados *Scopus*.

<b>País</b>	<b>Número total de registros na categoria de assunto <i>All Sciences (AS)</i> e frequência em porcentagem</b>	<b>Número total de registros na categoria de assunto <i>Social Sciences &amp; Humanities (SS&amp;H)</i> e frequência em porcentagem</b>	<b>Número total de registros nas categorias de assunto <i>All Sciences (AS)</i> e <i>Social Sciences &amp; Humanities (SS&amp;H)</i>, além de frequência em porcentagem</b>
<i>Australia</i>	7 (0,53%)	5 (0,69%)	12 (0,59%)
<i>Austria</i>	131 (9,98%)	24 (3,33%)	155 (7,62%)
<i>Belgium</i>	3 (0,23%)	4 (0,56%)	7 (0,34%)
<i>Botswana</i>	1 (0,08%)	1 (0,14%)	2 (0,1%)
<i>Brazil</i>	14 (1,07%)	9 (1,25%)	23 (1,13%)
<i>Canada</i>	5 (0,38%)	35 (4,86%)	40 (1,97%)
<i>China</i>	4 (0,3%)	3 (0,42%)	7 (0,34%)
<i>France</i>	1 (0,08%)	0 (0%)	1 (0,05%)
<i>Finland</i>	0 (0%)	1 (0,14%)	1 (0,05%)
<i>Germany</i>	113 (8,61%)	29 (4,03%)	142 (6,98%)
<i>India</i>	0 (0%)	7 (0,97%)	7 (0,34%)
<i>Iran</i>	2 (0,15%)	2 (0,28%)	4 (0,2%)
<i>Italy</i>	2 (0,15%)	1 (0,14%)	3 (0,15%)
<i>Mexico</i>	4 (0,3%)	7 (0,97%)	11 (0,54%)
<i>Netherlands</i>	70 (5,33%)	28 (3,89%)	98 (4,82%)
<i>Nigeria</i>	0 (0%)	2 (0,28%)	2 (0,1%)
<i>Portugal</i>	56 (4,27%)	11 (1,53%)	67 (3,3%)
<i>South Africa</i>	1 (0,08%)	1 (0,14%)	2 (0,1%)
<i>Spain</i>	1 (0,08%)	1 (0,14%)	2 (0,1%)
<i>Sudan</i>	1 (0,08%)	1 (0,14%)	2 (0,1%)
<i>Sweden</i>	0 (0%)	3 (0,42%)	3 (0,15%)
<i>Switzerland</i>	3 (0,23%)	0 (0%)	3 (0,15%)
<i>United Kingdom</i>	211 (16,07%)	140 (19,44%)	351 (17,27%)
<i>United States</i>	683 (52,02%)	405 (56,25%)	1088 (53,52%)
<b>TOTAL</b>	1313 (64,58% do número total de registros para ambas as categorias de assunto)	720 (35,42% do número total de registros para ambas as categorias de assunto)	2033 (100%)

**Fonte:** Elaborado pelos autores.

Nesta tabela no que diz respeito ao número total de registros recuperados apenas na categoria *Social Sciences & Humanities*, os países *United States* (405 registros, 56,25%), *United Kingdom* (140 registros, 19,44%) e *Canada* (35 registros, 4,86%) demonstraram as maiores quantidades e subsequentes frequências de ocorrência de registros. Ademais, o *Brazil* expôs poucos registros (9 registros, 1,25%), se comparado ao volume recuperado, por exemplo, para *Portugal* (11 registros, 1,53%). Isto reproduz muito ao que geralmente tem sido exposto em pesquisas sobre preservação digital: as comunidades envolvidas com o problema e que contribuem consideravelmente para o desenvolvimento de estudos científicos e de práticas em prol da preservação e do acesso utilizável, por longo prazo, ao conteúdo dos materiais digitais,

estão mais localizadas na região da Europa e da América do Norte e, em bem menor quantidade na América do Sul, onde ainda o tema é recente e carente de discussões teóricas e práticas.

### 3.3 Base de dados Web of Science<sup>3</sup>

Em referência ao levantamento executado nesta base de dados, obteve-se um conjunto expressivo de resultados (1141543 registros), sendo que deste volume 93,83% correspondeu à categoria *All Citation Databases* (1071155 registros) e 6,17% a categoria *Social Sciences Citation Index* (70388 registros). Assim sendo, a base de dados *Web of Science* se configurou no presente processo de levantamento, como sendo a fonte onde se recuperou o maior número de estudos produzidos no mundo sobre o tema. Logo, pode-se estabelecer na Tabela 6 a partir do número total de resultados obtidos nesta base para estas categorias, a distribuição ao longo de dez anos da quantidade e subsequente frequência da ocorrência dos registros relacionados aos assuntos de preservação digital.

**Tabela 6** - Distribuição da quantidade e frequência da ocorrência de registros obtidos para as categorias de assunto no período entre 2003-2013, base de dados *Web of Science*.

Ano	Número total de registros na categoria <i>All Citation Databases (ACD)</i> e frequência em porcentagem	Número total de registros na categoria <i>Social Sciences Citation Index (SSCI)</i> e frequência em porcentagem	Número total de registros nas categorias <i>All Citation Databases (ACD)</i> e <i>Social Sciences Citation Index (SSCI)</i> , além da frequência em porcentagem
2003	76517 (7,143%)	3800 (5,399%)	80317 (7,036%)
2004	81936 (7,649%)	3686 (5,237%)	85622 (7,501%)
2005	87748 (8,192%)	4750 (6,748%)	92498 (8,103%)
2006	95559 (8,921%)	4837 (6,872%)	100396 (8,795%)
2007	102848 (9,602%)	5819 (8,267%)	108667 (9,519%)
2008	109221 (10,197%)	6738 (9,573%)	115959 (10,158%)
2009	115624 (10,794%)	7679 (10,91%)	123303 (10,801%)
2010	112498 (10,502%)	8907 (12,654%)	121405 (10,635%)
2011	122120 (11,401%)	9964 (14,156%)	132084 (11,571%)
2012	125471 (11,714%)	10454 (14,852%)	135925 (11,907%)
*2013	41613 (3,885%)	3754 (5,333%)	45367 (3,974%)
<b>TOTAL</b>	1071155 (93,83% do número total de registros para ambas as categorias)	70388 (6,17% do número total de registros para ambas as categorias)	1141543 (100%)

**Fonte:** Elaborado pelos autores.

<sup>3</sup> Adotou-se o recurso *Search*, sendo que todas as expressões de busca foram inseridas no campo *Search* na opção *Topic*. Para cada expressão na opção de *Citation Databases*, selecionaram todas as bases (*Science Citation Index Expanded*, *Social Sciences Citation Index*, *Arts & Humanities Citation Index*, *Conference Proceedings Citation Index - Science* e *Conference Proceedings Citation Index - Social Science & Humanities*) numa primeira busca e, depois, apenas a base *Social Sciences Citation Index* na segunda busca. Em ambas as buscas desenvolvidas, escolheu-se no campo de *Limits* o valor *From 2003 to 2013* em *Timespan*.

Identificou-se por meio da tabela que no número total de registros obtidos apenas para a categoria de assunto *Social Sciences Citation Index*, o período entre os anos de 2009 a 2011 (26550 registros, 37,72%) exibiu um alto índice de produtividade de estudos durante estes dez anos. As maiores quantidades e frequências de ocorrência de registros se sucederam nos anos de 2012, 2011 e 2010 (29325 registros, 41,662%). Sendo assim, os resultados que foram recuperados na base *Web of Science* seguem uma tendência muito semelhante a já observada na base *Scopus*, onde o período entre 2009 a 2011 e os anos de 2010 e 2012 se destacam na produção de trabalhos ligados ao âmbito da preservação digital.

Além disto, pode-se elaborar na Tabela 7 com base no número total de resultados obtidos nesta base para as categorias *All Citation Databases* e *Social Sciences Citation Index*, a distribuição da quantidade e frequência da ocorrência de registros relacionados aos assuntos de preservação digital nos diversos tipos de documentos.

**Tabela 7** - Distribuição da quantidade e frequência da ocorrência de registros obtidos para os diversos tipos de documento nas categorias de assunto, base de dados *Web of Science*.

Número total de registros para os diversos tipos de documento na categoria <i>All Citation Databases (ACD)</i> e frequência em porcentagem		Número total de registros para os diversos tipos de documento na categoria <i>Social Sciences Citation Index (SSCI)</i> e frequência em porcentagem		Número total de registros para os diversos tipos de documento nas categorias <i>All Citation Databases (ACD)</i> e <i>Social Sciences Citation Index (SSCI)</i> , além de frequência em porcentagem	
<i>Art Exhibit Review</i>	8 (0,001 %)	<i>Art Exhibit Review</i>	0 (0%)	<i>Art Exhibit Review</i>	8 (0,001%)
<i>Article</i>	811344 (70,274%)	<i>Article</i>	55453 (76,029%)	<i>Article</i>	866797 (70,616%)
<i>Biographical Item</i>	72 (0,006%)	<i>Biographical Item</i>	16 (0,022%)	<i>Biographical Item</i>	88 (0,007%)
<i>Bibliography</i>	48 (0,004%)	<i>Bibliography</i>	20 (0,027%)	<i>Bibliography</i>	68 (0,006%)
<i>Book Chapter</i>	1355 (0,117%)	<i>Book Chapter</i>	52 (0,071%)	<i>Book Chapter</i>	1407 (0,115%)
<i>Book Review</i>	2124 (0,184%)	<i>Book Review</i>	1637 (2,244%)	<i>Book Review</i>	3761 (0,306%)
<i>Correction</i>	1531 (0,133%)	<i>Correction</i>	82 (0,112%)	<i>Correction</i>	1613 (0,131%)
<i>Database Review</i>	24 (0,002%)	<i>Database Review</i>	6 (0,008%)	<i>Database Review</i>	30 (0,002%)
<i>Editorial Material</i>	7615 (0,66%)	<i>Editorial Material</i>	1942 (2,663%)	<i>Editorial Material</i>	9557(0,779%)
<i>Excerpt</i>	4 (0%)	<i>Excerpt</i>	0 (0%)	<i>Excerpt</i>	4 (0%)
<i>Fiction Creative Prose</i>	4 (0%)	<i>Fiction Creative Prose</i>	0 (0%)	<i>Fiction Creative Prose</i>	4 (0%)
<i>Film Review</i>	4 (0%)	<i>Film Review</i>	0 (0%)	<i>Film Review</i>	4 (0%)
<i>Hardware Review</i>	4 (0%)	<i>Hardware Review</i>	0 (0%)	<i>Hardware Review</i>	4 (0%)
<i>Letter</i>	2622 (0,227%)	<i>Letter</i>	177 (0,243%)	<i>Letter</i>	2799 (0,228%)
<i>Meeting Abstract</i>	13105 (1,135%)	<i>Meeting Abstract</i>	1390 (1,906%)	<i>Meeting Abstract</i>	14495 (1,181%)
<i>Music Performance Review</i>	8 (0,001 %)	<i>Music Performance Review</i>	0 (0%)	<i>Music Performance Review</i>	8 (0,001%)
<i>News Item</i>	2042 (0,177%)	<i>News Item</i>	75 (0,103%)	<i>News Item</i>	2117 (0,172%)
<i>Proceedings Paper</i>	259192 (22,45%)	<i>Proceedings Paper</i>	2497 (3,424%)	<i>Proceedings Paper</i>	261689 (21,319%)
<i>Record Review</i>	52 (0,005%)	<i>Record Review</i>	0 (0%)	<i>Record Review</i>	52 (0,004%)
<i>Reprint</i>	152 (0,013%)	<i>Reprint</i>	24 (0,033%)	<i>Reprint</i>	176 (0,014%)
<i>Review</i>	53107 (4,6%)	<i>Review</i>	9538 (13,077%)	<i>Review</i>	62645 (5,104%)
<i>Script</i>	4 (0%)	<i>Script</i>	0 (0%)	<i>Script</i>	4 (0%)
<i>Software Review</i>	104 (0,009%)	<i>Software Review</i>	28 (0,038%)	<i>Software Review</i>	132 (0,011%)
<i>Theater Review</i>	8 (0,001 %)	<i>Theater Review</i>	0 (0%)	<i>Theater Review</i>	8 (0,001%)
<i>TV Review Radio Review</i>	8 (0,001 %)	<i>TV Review Radio Review</i>	0 (0%)	<i>TV Review Radio Review</i>	8 (0,001%)
<b>TOTAL</b>	1154541 (100%)	<b>TOTAL</b>	72937 (100%)	<b>TOTAL</b>	1227478 (100%)

Fonte: Elaborado pelos autores.

Nesta tabela constatou-se que em relação ao número total de registros obtidos somente na categoria de assunto *Social Sciences Citation Index*, os documentos denominados *Article*, *Review* e *Proceedings Paper* (67488 registros, 92,53%), apresentaram as maiores quantidades e subsequentes frequências de ocorrência de registros. Assim, seguindo o caminho traçado nos levantamentos realizados nas bases *ScienceDirect* e *Scopus*, os artigos publicados em revistas científicas e as revisões, no campo das Ciências Sociais, também se destacam na base *Web of Science* como grandes fontes de informação quanto a temática de preservação digital.

Pode-se também desenvolver na Tabela 8 a partir do número total de resultados obtidos nesta base para as categorias *All Citation Databases* e *Social Sciences Citation Index*, a distribuição da quantidade e frequência da ocorrência de registros relacionados a preservação digital no conjunto dos cinco primeiros países com maior volume de registros além do Brasil.

**Tabela 8** - Distribuição da quantidade e frequência da ocorrência de registros obtidos para os cinco primeiros países com maior volume de registros além do Brasil nas categorias de assunto, base de dados *Web of Science*.

<b>País</b>	<b>Número total de registros na categoria <i>All Citation Databases</i> (ACD) e frequência em porcentagem</b>	<b>Número total de registros na categoria <i>Social Sciences Citation Index</i> (SSCI) e frequência em porcentagem</b>	<b>Número total de registros nas categorias <i>All Citation Databases</i> (ACD) e <i>Social Sciences Citation Index</i> (SSCI), além de frequência em porcentagem</b>
<i>Argentina</i>	0 (0%)	1 (0%)	1 (0%)
<i>Australia</i>	16 (0%)	4418 (8,09%)	4434 (0,59%)
<i>Austria</i>	38 (0,01%)	0 (0%)	38 (0,01%)
<i>Belgium</i>	3 (0%)	1 (0%)	4 (0%)
<i>Brazil</i>	22303 (3,19%)	1039 (1,9%)	23342 (3,1%)
<i>Canada</i>	20 (0%)	4660 (8,53%)	4680 (0,62%)
<i>Cuba</i>	0 (0%)	1 (0%)	1 (0%)
<i>England</i>	59683 (8,54%)	8647 (15,83%)	68330 (9,07%)
<i>France</i>	14 (0%)	0 (0%)	14 (0%)
<i>Germany</i>	98750 (14,13%)	10 (0,02%)	98760 (13,11%)
<i>India</i>	7 (0%)	4 (0,01%)	11 (0%)
<i>Italy</i>	133 (0,02%)	3 (0,01%)	136 (0,02%)
<i>Japan</i>	82066 (11,74%)	0 (0%)	82066 (10,89%)
<i>Mexico</i>	1 (0%)	5 (0,01%)	6 (0%)
<i>Netherlands</i>	6 (0%)	12 (0,02%)	18 (0%)
<i>New Zealand</i>	0 (0%)	6 (0,01%)	6 (0%)
<i>Norway</i>	3 (0%)	0 (0%)	3 (0%)
<i>Peoples R China</i>	153300 (21,93%)	28 (0,05%)	153328 (20,35%)
<i>Portugal</i>	7743 (1,11%)	216 (0,4%)	7959 (1,06%)
<i>South Africa</i>	2 (0%)	5 (0,01%)	7 (0%)
<i>Spain</i>	5 (0%)	32 (0,06%)	37 (0%)
<i>Sudan</i>	1 (0%)	0 (0%)	1 (0%)
<i>Sweden</i>	2 (0%)	1 (0%)	3 (0%)
<i>Taiwan</i>	26 (0%)	2547 (4,66%)	2573 (0,34%)
<i>USA</i>	274819 (39,32%)	32977 (60,38%)	307796 (40,85%)
<b>TOTAL</b>	698941 (92,75% do número total de registros para ambas as categorias)	54613 (7,25% do número total de registros para ambas as categorias)	753554 (100%)

**Fonte:** Elaborado pelos autores.

Foi possível verificar através da tabela que no número total de registros obtidos somente na categoria de assunto *Social Sciences Citation Index*, os países *USA* (32977 registros, 60,38%), *England* (8647 registros, 15,83%) e *Canada* (4660 registros, 8,53%) possuíram as mais elevadas quantidades e subsequentes frequências de ocorrência de registros. Ademais, o *Brazil* expôs grandes quantidades e frequências de ocorrência de registros (1039 registros, 1,9%), se comparado ao volume de registros recuperados, por exemplo, para *Portugal* (216 registros, 0,4%). Apesar de não incluir o campo das Humanidades e abranger apenas o campo das Ciências Sociais, na base *Web of Science* de forma similar ao acontecido no levantamento da base *Scopus*, as regiões da Europa e da América do Norte se localizam as comunidades que mais cooperam e colaboram com a área da preservação digital.

### **3.4 Base de dados LISA<sup>4</sup>**

No levantamento elaborado nesta base de dados, obteve-se um conjunto relativamente expressivo de resultados (52171 registros), sendo que deste volume 99,66% equivaleram a Periódicos acadêmicos (51993 registros), 0,33% a Procedimentos e Trabalhos de Conferência (170 registros) e, por fim, 0,02% a Outras Fontes (8 registros). Tal como ocorreu nas bases *ScienceDirect*, *Scopus* e *Web of Science*, no cenário do campo das Ciências Sociais, as revistas científicas também se configuraram como principais fontes nesta base para a recuperação de informações sobre preservação digital às comunidades interessadas. Desta forma, a base de dados LISA se consolidou no presente processo de levantamento, como sendo a fonte onde se recuperou o segundo maior número estudos produzidos no mundo sobre a preservação digital. Isto posto, pode-se estabelecer na Tabela 9 a partir do número total de resultados obtidos nesta base, a distribuição ao longo de dez anos da quantidade e subsequente frequência da ocorrência dos registros associados aos assuntos de preservação digital.

---

<sup>4</sup> Adotaram-se o recurso de Busca avançada em Avançado, sendo que todas as expressões de busca foram inseridas na opção Assunto principal (todos) - SU. Para o campo Data de publicação escolheu-se a opção Intervalo de data específico e depois incluídos os valores entre os anos de 2003 e 2013. Não se utilizaram os filtros de Limitar a: Revisado por especialistas e Duplicatas: Incluir documentos duplicados. Ademais, foram selecionados todas as opções de fontes em Tipos de fontes, todas as opções de documentos em Tipos de documento e todos os idiomas disponíveis em Idiomas.

**Tabela 9** - Distribuição da quantidade e frequência da ocorrência de registros obtidos no período entre 2003-2013, base de dados LISA.

Ano	Número total de registros e frequência em porcentagem
2003	4400 (8,43%)
2004	4527 (8,68%)
2005	5790 (11,1%)
2006	6260 (12%)
2007	4397 (8,43%)
2008	4720 (9,05%)
2009	5994 (11,49%)
2010	5736 (10,99%)
2011	5446 (10,44%)
2012	4526 (8,67%)
*2013	384 (0,74%)
<b>TOTAL</b>	<b>52180 (100%)</b>

**Fonte:** Elaborado pelos autores.

Na tabela observou-se que o número total de resultados obtidos (52180 registros) foi superior ao anteriormente referido. O período entre os anos de 2009 a 2011 (17176 registros, 32,92%) exibiu um alto índice de produtividade de trabalhos durante estes dez anos, enquanto que os anos de 2006, 2009 e 2005 (18044 registros, 34,59%) demonstraram as maiores quantidades e frequências de ocorrência de registros. Os resultados obtidos na base LISA também refletiram a tendência presente nas bases *Scopus* e *Web of Science*, em especial a grande produção de estudos sobre preservação digital no ano de 2009 e no período entre 2009 a 2011. Assim, através deste conjunto de resultados advindos destas bases, as comunidades interessadas e as áreas que integram o campo da Ciência da Informação como, por exemplo, a Biblioteconomia e a Arquivologia, podem se pautar nestes anos, períodos e fontes de informação identificadas para o desenvolvimento de suas futuras pesquisas e práticas em prol da preservação e acesso, a longo prazo, ao conteúdo dos recursos digitais.

#### **4 RESULTADOS OBTIDOS NAS BASES DE DADOS E REPOSITÓRIOS COM ABRANGÊNCIA EM ESTUDOS CIENTÍFICOS NACIONAIS**

##### **4.1 Biblioteca Digital Brasileira de Teses e Dissertações do IBICT<sup>5</sup>**

Em relação ao levantamento efetuado neste repositório, obteve-se um conjunto baixo de resultados (8 registros). Desta maneira, a BDBTD do IBICT se instituiu no presente processo, como sendo a quinta maior fonte na obtenção de estudos nacionais sobre preservação digital.

<sup>5</sup> Adotou-se o recurso de Procura avançada, sendo que todas as expressões de busca foram inseridas no campo Assunto na opção Todas as palavras. Para cada expressão no filtro Grau, empregou-se a opção Dissertações numa primeira busca e depois a opção Teses na segunda busca. Não se utilizaram os campos Autor, Resumo, Título, Contribuidor e Instituição de Defesa. Os filtros denominados País e Idioma também não foram usados. Além disso, no filtro Ano de defesa foram inseridos os valores 2003 em A partir de e 2013 em e/ou até.

Sendo assim, pode-se estabelecer na Tabela 10 a partir do número total de resultados obtidos neste repositório, a distribuição ao longo de oito anos da quantidade e subsequente frequência da ocorrência de dissertações e teses relacionadas aos assuntos de preservação digital.

**Tabela 10** - Distribuição da quantidade e frequência da ocorrência de dissertações e teses obtidas no período entre 2004-2012, BDBTD do IBICT.

Ano	Número de Dissertações e frequência em porcentagem	Número de Teses e frequência em porcentagem
2004	1 (12,5%)	0 (0%)
2005	0 (0%)	0 (0%)
2006	0 (0%)	0 (0%)
2007	0 (0%)	0 (0%)
2008	1 (12,5%)	1 (12,5%)
2009	0 (0%)	0 (0%)
2010	1 (12,5%)	0 (0%)
2011	0 (0%)	0 (0%)
2012	1 (12,5%)	0 (0%)
<b>TOTAL</b>	4 (50%)	1 (12,5%)

**Fonte:** Elaborado pelos autores.

Nesta tabela constatou-se que apesar de existir um número total de dissertações e teses recuperadas (8 registros), com o emprego dos filtros para a obtenção de apenas dissertações (4 registros) e depois somente teses (1 registro), a soma dos resultados obtidos expostos na tabela (5 registros) não significou o número total recuperado na busca geral sem a aplicação destes recursos. Esta quantidade é compreensível se comparado com o período considerado normal para a produção de alguns tipos de documentos como, por exemplo, de artigos científicos, mas não pode ser visto como uma justificativa pontual para a carência de trabalhos na área de preservação digital. Além disso, no número total de registros obtidos, 50% correspondeu a dissertações com frequência de ocorrência de registros de 12,5% (1 registro) em 2004, 2008, 2010 e 2012. Ademais, 12,5% equivaleram a teses com incidência de 12,5% (1 registro) em 2008 e, por último, 37,5% referentes aos resultados não obtidos em decorrência da aplicação dos filtros. Estes anos identificados, podem significar uma tendência de que as dissertações e teses defendidas nos programas de pós-graduação do país, passaram a partir do ano de 2004 e de 2008, a dedicar-se ou dar maior atenção aos problemas e os demais aspectos envolvidos com processo de preservação digital.

## 4.2 Biblioteca Digital de Teses e Dissertações da USP<sup>6</sup>

Em contradição ao levantamento desenvolvido na BDBTD do IBICT, na BDTD da USP obteve-se um conjunto consideravelmente grande de resultados (517 registros) sem qualquer limitação de período de tempo. Deste modo, a BDTD da USP se caracterizou no presente processo de levantamento, como sendo a segunda maior fonte na recuperação de estudos nacionais sobre o tema. Por consequência, pode-se elaborar na Tabela 11 a partir do número total de resultados obtidos neste repositório, a distribuição ao longo de dez anos da quantidade e subsequente frequência da ocorrência dos registros para dissertações e teses associadas aos assuntos de preservação digital.

**Tabela 11** - Distribuição da quantidade e frequência da ocorrência de dissertações e teses obtidas no período entre 2003-2012, BDTD da USP.

Ano	Número de registros para Dissertações e frequência em porcentagem	Número de registros para Teses e frequência em porcentagem	Número total de registros e frequência em porcentagem
2003	29 (10,507%)	4 (1,99%)	33 (6,9%)
2004	5 (1,812%)	24 (11,94%)	29 (6,1%)
2005	0 (0%)	12 (5,97%)	12(2,5%)
2006	24 (8,696%)	17 (8,458%)	41 (8,6%)
2007	54 (19,565%)	21 (10,448%)	75 (15,7%)
2008	24 (8,696%)	28 (13,93%)	52 (10,9%)
2009	48 (17,391%)	21 (10,448%)	69 (14,5%)
2010	17 (6,159%)	11 (5,473%)	28 (5,9%)
2011	35 (12,681)	24 (11,94%)	59 (12,4%)
2012	28 (10,145%)	39 (19,403%)	67 (14%)
*2013	12 (4,348%)	0 (0%)	12 (2,5%)
<b>TOTAL</b>	276 (57,862% do número total de registros)	201 (42,138% do número total de registros)	477 (100%)

**Fonte:** Elaborado pelos autores.

Identificou-se por meio da tabela que no número total de resultados obtidos para dissertações, os anos de 2007, 2009 e 2011 (137 registros, 49,637%) possuíram elevadas quantidades e frequências de incidência de registros. Quanto ao número total de registros para teses, os anos de 2012, 2008, 2004 e 2011 (115 registros, 57,213%) se destacaram com altas

<sup>6</sup> Adotaram-se os recursos Pesquisa e posteriormente Avançada, sendo que todas as expressões de busca foram inseridas nos campos presentes em Pesquisa Avançada na opção em Palavras-chave. Para as expressões sem operadores booleanos aplicou-se o operador “e” disponível, sendo que os termos destas expressões assim como para aquelas com operadores foram inseridos em campos separados com a escolha para ambos da opção em Palavras-chave. Para as expressões com operadores utilizaram-se alguns operadores disponíveis (“e” para AND e “ou” para OR, p. ex.). No recurso Pesquisa, não se identificou uma opção para a escolha do período a ser abrangido na recuperação dos estudos. No filtro de tipos de documentos a serem obtidos foi selecionado somente a opção Todos documentos, com a obtenção de resultados houve o cálculo do número total de dissertações e teses recuperadas em cada expressão e, depois, somou-se os registros presentes no período entre os anos de 2003 a 2013. Para expressões com valores inferiores a cinco resultados, através do acesso ao conteúdo integral dos trabalhos realizou-se a análise de cada registro a fim de averiguar e coletar dados fornecidos sobre o grau, o ano de defesa, a unidade da instituição e a área de conhecimento relacionada com o trabalho.

quantidades e frequência de registros. Desta maneira, assim como foi observado no levantamento realizado na BDBTD do IBICT, no contexto das teses de doutorado, os anos de 2004 e 2008 também são constatados na BDTD da USP.

### 4.3 Biblioteca Digital da UNICAMP<sup>7</sup>

No que concerne ao levantamento desenvolvido na BD da UNICAMP, obteve-se um baixo volume de resultados (8 registros). Assim, semelhante ao levantamento da BDBTD do IBICT, a BD da UNICAMP constituiu-se no presente processo de levantamento como sendo também a quinta maior fonte na recuperação de estudos nacionais sobre preservação digital. Logo, pode-se estabelecer na Tabela 12 a partir do número total de resultados obtidos neste repositório, a distribuição ao longo de quatro anos da quantidade e frequência da ocorrência dos registros ligadas aos assuntos de preservação digital.

**Tabela 12** - Distribuição da quantidade e frequência da ocorrência de registros obtidos para a categoria Eventos no período entre 2004-2007, BD da UNICAMP.

Ano	Número de registros na categoria Eventos e frequência em porcentagem	Título do Evento
2004	7 (87,5%)	II Simpósio Internacional de Bibliotecas Digitais, 2004
2005	0 (0%)	-
2006	0 (0%)	-
2007	1 (12,5%)	Seminário Internacional de Bibliotecas Digitais Brasil SIBDB 2007
<b>TOTAL</b>	<b>8 (100%)</b>	

Fonte: Elaborado pelos autores.

Através da tabela foi possível observar que da quantidade de registros recuperados, apenas 12,5% correspondeu ao evento intitulado Seminário Internacional de Bibliotecas Digitais Brasil SIBDB 2007 e 87,5% ao denominado II Simpósio Internacional de Bibliotecas Digitais (2004). Assim, podemos compreender que estes eventos refletem nas mais recentes e importantes iniciativas criadas no contexto científico nacional para o âmbito da preservação

<sup>7</sup> Adotou-se o recurso Pesquisar, sendo que todas as expressões de busca foram inseridas no campo Pesquisar na opção Palavras-chave. Em Pesquisar não se observou uma opção para a escolha de diferentes operadores de busca booleanos para cada termo presente em cada expressão. Assim, selecionamos a opção Todas as palavras (AND) ao invés de Quaisquer palavras (OR), independente das expressões apresentarem ou não uso simultâneo de operadores distintos. No filtro Pesquisar em foi selecionado o item Todos dentro do grupo de opções disponíveis (Dissertações e Teses, Trabalhos de Conclusão de Curso, Eventos, Hemeroteca – CMU, Revistas Eletrônicas, Produção Técnico-Científica Digital e Coleções Especiais) e em Formato de Apresentação aplicou-se a opção com resumo. Não foi utilizado o campo Exibir número de resultados por página e o recurso Pesquisa por Index. Ademais, não se identificou no recurso Pesquisar uma opção para a escolha do período a ser abrangido na recuperação dos registros.

digital, sendo os anos de 2004 e 2007, onde aparecem nos levantamentos da BDBTD do IBICT e da BDTD da USP, se destacaram também na produção de trabalhos neste domínio.

#### 4.4 Biblioteca Digital de Teses e Dissertações da UNESP<sup>8</sup>

Ao contrário do levantamento executado na BD da UNICAMP, na BDTD da UNESP obteve-se um conjunto elevado de resultados (244 registros). Portanto, a BDTD da UNESP se apresentou no presente processo de levantamento como sendo a terceira maior fonte na obtenção de estudos nacionais sobre a temática tratada. Deste modo, pode-se estabelecer na Tabela 13 a partir do número total de resultados obtidos neste repositório para as opções de Assunto e Título, a distribuição ao longo de dez anos da quantidade e subsequente frequência da ocorrência de dissertações e teses correlatadas aos assuntos de preservação digital.

**Tabela 13** - Distribuição da quantidade e frequência da ocorrência de dissertações e teses obtidas para as opções de busca no período entre 2003-2013, BDTD da UNESP.

Ano	Número total de registros para Dissertações nas opções Assunto e Título, além de frequência em porcentagem	Número total de registros para Teses nas opções Assunto e Título, além de frequência em porcentagem	Número total de registros para Dissertações e Teses nas opções Assunto e Título, além de frequência em porcentagem
2003	0 (0%)	12 (5,02%)	12 (4,92%)
2004	1 (20%)	8 (3,35%)	9 (3,69%)
2005	1 (20%)	16 (6,69%)	17 (6,97%)
2006	0 (0%)	4 (1,67%)	4 (1,64%)
2007	0 (0%)	8 (3,35%)	8 (3,28%)
2008	0 (0%)	28 (11,72%)	28 (11,48%)
2009	1 (20%)	17 (7,11%)	18 (7,38%)
2010	0 (0%)	9 (3,77%)	9 (3,69%)
2011	1 (20%)	49 (20,5%)	50 (20,49%)
2012	1 (20%)	60 (25,1%)	61 (25%)
*2013	0 (0%)	28 (11,72%)	28 (11,48%)
<b>TOTAL</b>	5 (2,05% do número total de registros para dissertações e teses)	239 (97,95% do número total de registros para dissertações e teses)	244 (100%)

Fonte: Elaborado pelos autores.

<sup>8</sup> Adotaram-se os recursos de Busca Simples/Avançada e depois o de Busca Avançada, sendo que todas as expressões de busca foram inseridas no campo Palavras. Para cada expressão no filtro Campo empregou-se a opção Assunto numa primeira busca e depois a opção Título na segunda busca. Em ambas as buscas escolheu-se a opção Sim no recurso denominado Palavras adjacentes. No recurso Base selecionou-se a opção Biblioteca Digital de Teses e Dissertações. Em Filtros de busca aplicou-se a opção Todos para as opções de Idioma e Formato, sendo que em Ano de publicação de foi inserido o valor 2003 e em seguida o valor 2013 na opção até. Com a recuperação de resultados, houve o cálculo do número total de dissertações e teses obtidas nas buscas por Assunto e Título em cada expressão e, posteriormente, também se somou os registros presentes no período de 2003 a 2013 nas buscas por Assunto e Título em cada expressão. Para expressões com valores inferiores a cinco resultados, através do acesso ao conteúdo integral dos trabalhos houve uma análise de cada registro a fim de averiguar e coletar dados fornecidos sobre o grau, o ano de defesa, o programa de pós-graduação ou curso, a unidade da instituição além do título e a área correspondente.

Nesta tabela verificamos que no número total de registros recuperados para teses, os anos de 2012, 2011, 2008 e 2013 (165 registros, 69,04%) possuíram altas quantidades e frequências de incidência de registros. Contudo, quanto ao número total de registros para dissertações, houve poucas variações nas baixas quantidades e frequências de registros identificadas no decorrer do período (entre 0 e 1 registro, 0% e 20%). Os anos de 2008, 2011 e 2012 também são identificados, tal como observamos na BDBTD do IBICT e na BDTD da USP, em relação às teses defendidas nos programas de pós-graduação. Estes anos identificados são importantes para as comunidades interessadas quanto a obtenção de informações sobre preservação digital, com base na consulta de dissertações de mestrado e teses de doutorado.

#### 4.5 SciELO<sup>9</sup>

Em oposição ao levantamento executado na BDBTD do IBICT e na BD da UNICAMP, nesta base recuperou-se um número maior de resultados (28 registros). Por conseguinte, a SciELO constituiu-se no presente processo de levantamento como sendo a quarta maior fonte na recuperação de estudos nacionais sobre preservação digital. Assim, pode-se estabelecer na Tabela 14 a partir do número total de resultados recuperados nesta base, um *ranking* dos periódicos nacionais de diversas áreas com maior número e frequência de ocorrência de registros para os assuntos relacionados a preservação digital.

**Tabela 14** - Distribuição da quantidade e frequência da ocorrência de registros obtidos para os diversos periódicos nacionais de diferentes áreas, SciELO.

Posição no ranking e título do periódico	Número total de registros	Frequência em porcentagem do número total de registros
1. Ciência da Informação	11	39,286%
2. Revista Árvore	3	10,714%
3. Revista de Antropologia	2	7,143%
4. <i>Brazilian Journal of Oceanography</i>	2	7,143%
5. Ciência Rural	2	7,143%
6. Engenharia Agrícola	2	7,143%
7. Anais da Academia Brasileira de Ciências	1	3,571%
8. Fisioterapia em Movimento	1	3,571%
9. Revista Brasileira de Coloproctologia	1	3,571%
10. Revista Ceres	1	3,571%
11. Revista Brasileira de Engenharia Agrícola e Ambiental	1	3,571%
12. Revista de Odontologia da UNESP	1	3,571%
<b>TOTAL</b>	<b>28</b>	<b>100%</b>

**Fonte:** Elaborado pelos autores.

<sup>9</sup> Adotou-se a opção Pesquisa artigos, sendo que todas as expressões de busca foram inseridas no campo Entre com uma ou mais palavras. Para cada expressão utilizou-se o campo Método na opção integrada e o campo onde na opção Brasil. Com a recuperação de resultados, escolheu-se a opção Resumo no campo de índices. Durante o processo não se identificou uma opção para a seleção do período a ser abrangido na recuperação dos estudos presentes na base de dados.

A partir da tabela observou-se que a grande maioria dos periódicos identificados com exceção do primeiro, apresentou uma quantidade e frequência de ocorrência de registros bem aproximada. Além disso, o periódico denominado Ciência da Informação (11 registros, 39,286%) que abrange o campo da Ciência da Informação e da informação em geral tal como o periódico Revista Árvore (3 registros, 10,714%) que compreende o campo da Ciência Florestal e áreas afins, se destacaram no presente levantamento com as maiores quantidades e frequências de ocorrência de registros. Percebe-se então que a preservação digital é um tema de pesquisa multidisciplinar, pois não é alvo de investigação de somente as áreas que lidam com a gestão da informação, mas também aquelas com origens nos campos das Ciências Biológicas e Exatas. Logo, as organizações envolvidas com tal temática podem então adotar como estratégia de pesquisa para obtenção de informações a não limitação a áreas específicas.

Ademais, pode-se demonstrar na Tabela 15 através do número total de resultados obtidos nesta base, a distribuição ao longo de dezesseis anos da quantidade e subseqüente frequência da ocorrência de registros ligados aos assuntos de preservação digital.

**Tabela 15** - Distribuição da quantidade e frequência da ocorrência de registros obtidos no período entre 1997-2013, SciELO.

Ano	Número total de registros	Frequência em porcentagem do número total de registros
1997	2	7,143%
1998	0	0%
1999	1	3,571%
2000	0	0%
2001	1	3,571%
2002	2	7,143%
2003	0	0%
2004	4	14,286%
2005	2	7,143%
2006	2	7,143%
2007	0	0%
2008	2	7,143%
2009	3	10,714%
2010	3	10,714%
2011	3	10,714%
2012	2	7,143%
*2013	1	3,571%
<b>TOTAL</b>	<b>28</b>	<b>100%</b>

**Fonte:** Elaborado pelos autores.

Nesta tabela também se constatou um volume aproximado de estudos produzidos ao longo destes dezesseis anos, com exceção do ano de 2004 (4 registros, 14,286%) onde ocorreu a maior quantidade e subseqüente frequência de incidência de registros. Os períodos de maior

produtividade observados na tabela foram entre os anos de 2009 a 2011 (9 registros, 32,142%), em seguida entre 2003 a 2005 (6 registros, 21,429%) e, por último, entre os anos de 2006 a 2008 (4 registros, 14,286%). Desta maneira, o ano de 2004 se destaca na produção de artigos científicos juntamente com o período de 2009 a 2011, onde podemos observar uma produção ainda singela, mas crescente e constante de estudos sobre a preservação digital.

#### 4.6 Brapci<sup>10</sup>

No levantamento realizado nesta base de dados, obteve-se um conjunto elevado de resultados (681 registros). Assim sendo, a Brapci se manifestou no presente processo de levantamento, como sendo a fonte onde se recuperou o maior número de estudos nacionais sobre o tema. Desta maneira, pode-se estabelecer na Tabela 16 a partir do número total de resultados obtidos nesta base, um *ranking* dos periódicos nacionais impressos e eletrônicos que compreendem o campo da Ciência da Informação e, que possuem o maior número e frequência de ocorrência de registros para os assuntos relacionados à preservação digital.

**Tabela 16** - Distribuição da quantidade e frequência da ocorrência de registros obtidos para os diversos periódicos nacionais do campo de Ciência da Informação, base de dados Brapci.

Posição no <i>ranking</i> e título do periódico	Número total de registros	Frequência em porcentagem do número total de registros
1. Ciência da Informação	86	12,63%
2. Informação & Sociedade: Estudos	72	10,57%
3. Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação	59	8,66%
4. DataGramaZero	58	8,52%
5. Perspectivas em Ciência da Informação	51	7,49%
6. Arquivística.net	47	6,9%
7. Revista ACB: Biblioteconomia em Santa Catarina	35	5,14%
8. Transinformação	34	4,99%
9. Ponto de Acesso	33	4,85%
10. Informação & Informação	27	3,96%
11. Revista Brasileira de Biblioteconomia e Documentação	26	3,82%
12. Biblionline	26	3,82%
13. Revista Digital de Biblioteconomia & Ciência da Informação	24	3,52%
14. Liinc em revista	18	2,64%
15. InCID: Revista de Ciência da Informação e Documentação	17	2,5%
16. BIBLOS - Revista do Instituto de Ciências Humanas e da Informação	17	2,5%

<sup>10</sup> Adotou-se o recurso de Busca para a inserção no campo Palavras-chave de todas as expressões de busca. Para a escolha do período a ser abrangido na recuperação dos estudos presentes na base, empregou-se a opção Definir período e depois se determinou os valores entre os anos de 2003 até 2013. Com a ocorrência de resultados, foram selecionados todos os “quadrados” referentes a estes resultados obtidos e em seguida se utilizou o recurso Seleção para a análise dos resultados e posteriormente a produção das respectivas tabelas.

17. Em Questão: Revista da Faculdade de Biblioteconomia e Comunicação da UFRGS	16	2,35%
18. ETD - Educação Temática Digital	15	2,2%
19. Inclusão Social	7	1,03%
20. Comunicação & Informação	4	0,59%
21. Perspectivas em Gestão & Conhecimento	4	0,59%
22. Revista Ibero-Americana de Ciência da Informação	2	0,29%
23. Tendências da Pesquisa Brasileira em Ciência da Informação	2	0,29%
24. <i>Brazilian Journal of Information Science</i>	1	0,15%
<b>TOTAL</b>	<b>681</b>	<b>100%</b>

**Fonte:** Elaborado pelos autores.

Na tabela constatou-se que um grupo de periódicos demonstrou uma quantidade e frequência de ocorrência de registros um pouco aproximada, sendo que em alguns periódicos o número total de registros recuperados foi o mesmo. Foi possível verificar também que os periódicos intitulados Ciência da Informação (86 registros, 12,63%), Informação & Sociedade: Estudos (72 registros, 10,57%) e Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação (59 registros, 8,66%), possuíram as maiores quantidades e subsequentes frequências de ocorrência de registros. Estas revistas, em especial o periódico Ciência da Informação por sua também identificação no levantamento realizado na SciELO, podem ser apontadas como grandes fontes de informação para o campo das Ciência da Informação e para as comunidades interessadas na obtenção de artigos nacionais sobre preservação digital.

Além disso, pode-se apresentar na Tabela 17 por meio do número total de resultados obtidos nesta base, a distribuição ao longo de dezesseis anos da quantidade e subseqüente frequência da ocorrência de registros relacionados aos assuntos de preservação digital.

**Tabela 17** - Distribuição da quantidade e frequência da ocorrência de registros obtidos no período entre 1997-2013, base de dados Brapci.

Ano	Número total de registros	Frequência em porcentagem do número total de registros
1997	1	0,15%
1998	3	0,44%
2000	1	0,15%
2001	2	0,29%
2002	2	0,29%
2003	3	0,44%
2004	50	7,34%
2005	51	7,49%
2006	101	14,83%
2007	104	15,27%
2008	86	12,63%
2009	65	9,54%
2010	102	14,98%
2011	73	10,72%
2012	36	5,29%
*2013	1	0,15%
<b>TOTAL</b>	<b>681</b>	<b>100%</b>

**Fonte:** Elaborado pelos autores.

Em contrapartida ao reduzido volume de estudos produzidos nos seis primeiros anos anteriores, através da tabela foi possível identificar um crescimento considerável a partir do ano de 2004 (50 registros, 7,34%). Os períodos de maior produtividade identificados na tabela foram entre os anos de 2007 a 2009 (255 registros, 37,44%), em seguida entre 2010 a 2012 (211 registros, 30,99%), depois entre 2004 a 2006 (202 registros, 29,66%) e, por último, entre anos de 2001 a 2003 (7 registros, 1,02%). Assim como no levantamento realizado na SciELO, o ano de 2004 na base Brapci se apresenta como aquele onde a temática de preservação digital passa ser objeto de investigações pelos pesquisadores em seus artigos. Embora se tratando apenas de trabalhos na área da Ciência da Informação, entre o período de 2007 a 2009, se constatou também um crescimento progressivo de publicações sobre tal temática, abrangendo mesmo que, parcialmente, os períodos identificados na SciELO.

## 5 CONSIDERAÇÕES FINAIS

A partir dos resultados obtidos no levantamento presente neste trabalho, no panorama científico internacional, a *Web of Science*, *LISA*, *Scopus* e *ScienceDirect*, configuraram-se, nesta ordem, como sendo as bases onde se recuperou os maiores números de estudos produzidos no mundo sobre preservação digital. Nestas bases, no contexto do campo das Ciências Sociais, o período entre os anos de 2009 a 2011 e o ano de 2012 se destacaram na produção estudos sobre o assunto, sendo que as revistas científicas e as revisões que abrangem as áreas que compõem este campo se sobressaíram como principais fontes para a obtenção de informações. As comunidades presentes na Europa (Inglaterra) e na América do Norte (Estados Unidos e Canadá), envolvidas com problemática da preservação e do acesso de longo prazo, são as que mais colaboram no desenvolvimento de estudos e iniciativas neste domínio.

Além disso, no cenário científico nacional, a Brapci, seguida da BDTD da USP, da BDTD da UNESP, da SciELO e, por último, da BDBTD do IBICT assim como da BD da UNICAMP, se constituíram como sendo as bases e os repositórios onde se obteve os maiores números de estudos nacionais sobre o tema. No âmbito das dissertações e teses defendidas nos programas de pós-graduação do país, a partir dos anos de 2004 e de 2008, os pesquisadores passaram a dedicar-se aos problemas e demais aspectos envolvidos com o assunto, enquanto que para os eventos científicos nacionais, os anos de 2004 e 2007, se destacaram também na produção de trabalhos. No contexto dos artigos científicos da área da Ciência da Informação, o periódico *Ciência da Informação*, o ano de 2004 e o período de 2007 a 2009 se consolidaram como grandes fontes para a obtenção de estudos.

Apesar da grande quantidade de resultados recuperados nas diversas bases e repositórios levantados, se verificou que o assunto ainda se caracteriza por ser novo e carente de iniciativas com relação ao contexto científico nacional. Poucos estudos desenvolvidos foram encontrados, por exemplo, na BDBTD do IBICT, na BD da UNICAMP e na SciELO. Portanto, torna-se fundamental a realização de mais pesquisas sobre esta temática em benefício da comunidade científica nacional e internacional além da sociedade. A identificação deste conjunto de dados nestas bases e repositórios poderá auxiliar as comunidades interessadas e as áreas que integram o campo das Ciências Sociais tal como, por exemplo, a Ciência da Informação, no desenvolvimento de suas futuras pesquisas e práticas em prol da preservação e acesso, à longo prazo, aos conteúdos digitais.

## REFERÊNCIAS

BASE DE DADOS REFERENCIAL DE ARTIGOS DE PERIÓDICOS EM CIÊNCIA DA INFORMAÇÃO (BRAPCI). Disponível em: <<http://www.brapci.ufpr.br/>>. Acesso em: 27 fev. 2014.

CADERNOS DE BIBLIOTECONOMIA, ARQUIVÍSTICA E DOCUMENTAÇÃO (CADERNOS BAD). Disponível em: <[http://www.apbad.pt/Edicoes/Edicoes\\_Cadernos.htm](http://www.apbad.pt/Edicoes/Edicoes_Cadernos.htm)>. Acesso em: 27 fev. 2014.

COORDENAÇÃO DE APERFEICOAMENTO DE PESSOAL DE NÍVEL SUPERIOR (CAPES). **Portal de Periódicos**. Disponível em: <<http://www.periodicos.capes.gov.br/>>. Acesso em: 27 fev. 2014.

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA (IBICT). **Biblioteca Digital Brasileira de Teses e Dissertações**. Disponível em: <<http://btdt.ibict.br/>>. Acesso em: 27 fev. 2014.

LIBRARY AND INFORMATION SCIENCE ABSTRACTS (LISA). Disponível em: <<http://search.proquest.ez31.periodicos.capes.gov.br/lisa/index>>. Acesso em: 27 fev. 2014.

RUSSEL, Kelly; SERGEANT, Derek. The Cedars project: implementing a model for distributed digital archives. **RLG DigiNews**, n. 3, v. 3, jun. 1999. Disponível em: <<http://worldcat.org/arcviewer/2/OCC/2009/07/29/H1248898656915/viewer/file2.html>>. Acesso em: 27 fev. 2014.

SCIENCEDIRECT. Disponível em: <<http://www-sciencedirect-com.ez31.periodicos.capes.gov.br/>>. Acesso em: 27 fev. 2014.

SCIENTIFIC ELECTRONIC LIBRARY ONLINE (SCIELO). Disponível em: <<http://www.scielo.org/php/index.php>>. Acesso em: 27 fev. 2014.

SCOPUS. Disponível em: <<http://www-scopus-com.ez31.periodicos.capes.gov.br/home.url>>. Acesso em: 27 fev. 2014.

UNIVERSIDADE DE SÃO PAULO (USP). **Biblioteca Digital de Teses e Dissertações**. Disponível em: <<http://www.teses.usp.br/>>. Acesso em: 27 fev. 2014.

UNIVERSIDADE ESTADUAL DE CAMPINAS (UNICAMP). **Biblioteca Digital**. Disponível em: <<http://www.bibliotecadigital.unicamp.br/>>. Acesso em: 27 fev. 2014.

UNIVERSIDADE ESTADUAL PAULISTA JÚLIO DE MESQUITA FILHO (UNESP). **Biblioteca Digital de Teses e Dissertações**. Disponível em: <<http://www.unesp.br/portal#!/cgb/bibliotecas-digitais/cthedra-biblioteca-digital-teses/>>. Acesso em: 27 fev. 2014.

WEOFSCIENCE. Disponível em: <[http://apps.webofknowledge.com.ez31.periodicos.capes.gov.br/WOS/GeneralSearch\\_input.do?product=WOS&search\\_mode=GeneralSearch&SID=3DNI7L4GiboaLBM1jeA&preferencesSaved=>](http://apps.webofknowledge.com.ez31.periodicos.capes.gov.br/WOS/GeneralSearch_input.do?product=WOS&search_mode=GeneralSearch&SID=3DNI7L4GiboaLBM1jeA&preferencesSaved=>)>. Acesso em: 27 fev. 2014.