

Ten-year Analysis of University of Minho Green OA Self-Archiving Mandate

Yassine Gargouri, Vincent Larivière, Stevan Harnad

Abstract: *University of Minho adopted the first university-wide Green Open Access Mandate in Europe in 2004, requiring all research output to be self-archived in the institution's repository. The mandate was upgraded in 2011 to designate the repository as the sole mechanism for submitting publications for individual and institutional research performance assessment. A 10-year analysis shows that deposit rates are increasing and deposit delays are decreasing. Once the rest of the world follows Minho's example, universal Green OA will not be far behind.*

1. The 2004 Minho Mandate. In 2004, University of [Minho](#) in Portugal became the second university in the world, and the first in Europe, to adopt a university-wide [Green Open Access \(OA\) self-archiving mandate](#), requiring all researchers to deposit their publications in Minho's institutional repository, [RepositòriUM](#)¹. Minho's mandate was very successful: Within two years it generated deposit rates three times as high as those of non-mandating institutions (Gargouri et al 2010, **Figure 1**); it was also very influential, helping to spread and accelerate the adoption of OA mandates across Europe and worldwide.

2. The 2011 Liège Upgrade. Initially, a small financial reward was offered by Minho's Rector as an incentive for depositing, but after a few years this was no longer thought to be necessary. Nevertheless, it was noted that although Minho's deposit rates remained higher than non-mandatory rates, they had declined somewhat across the years, and were still far from 100%. So in 2011, the Minho mandate was upgraded to what has been found to be the strongest and most effective Green OA mandate (Gargouri et al 2012a) -- the Liège model mandate (Rentier & Thirion 2011):

(1) The deposit must be made *immediately* upon acceptance for publication, not after a delay, although access to the deposit can be set to *restricted access* instead of *public access* if the author elects to comply

¹ The first Green OA mandate was adopted by the [School of Electronics and Computer Science](#) at the University of Southampton in the UK in 2003, and the first university-wide Green OA mandate was adopted by [Queensland University of Technology](#) in the same year.

with a publisher OA embargo; for restricted access deposits, during the embargo the repository has a [Button](#) that allows users to request and authors to provide an individual copy with one click each (Sale et al. 2012).

(2) Deposit in the institutional repository is officially designated as the [sole mechanism](#) for submitting publications for [research performance review](#) or [funding](#).

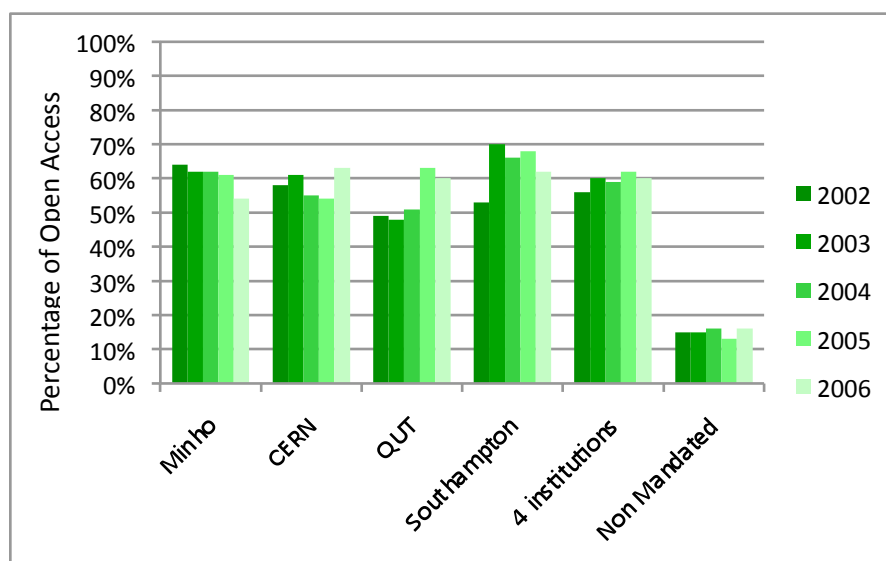


Figure 1. OA rates for 2002-2006 articles from Minho and 3 other institutions with Green OA self-archiving mandates, compared with non-mandated institutions. (Note that the Minho mandate was adopted in 2004, but the resulting deposits often include retroactive deposit of older papers as well. These robot-based OA rates, however, also include Minho articles made OA elsewhere on the Web, not just in Minho’s repository, hence they are higher than repository rates.) (Data are from [Gargouri et al 2010](#).)

We analyze here Minho’s deposit rate for OA’s principal target content -- peer-reviewed journal articles -- comparing Minho’s repository with three other repositories: Liege’s [ORBi](#), Surrey’s [SRI](#) (mandated), Lancaster’s [EPrints](#) (non-mandated), as well as the average for 26 mandated UK repositories and 73 non-mandated UK repositories.

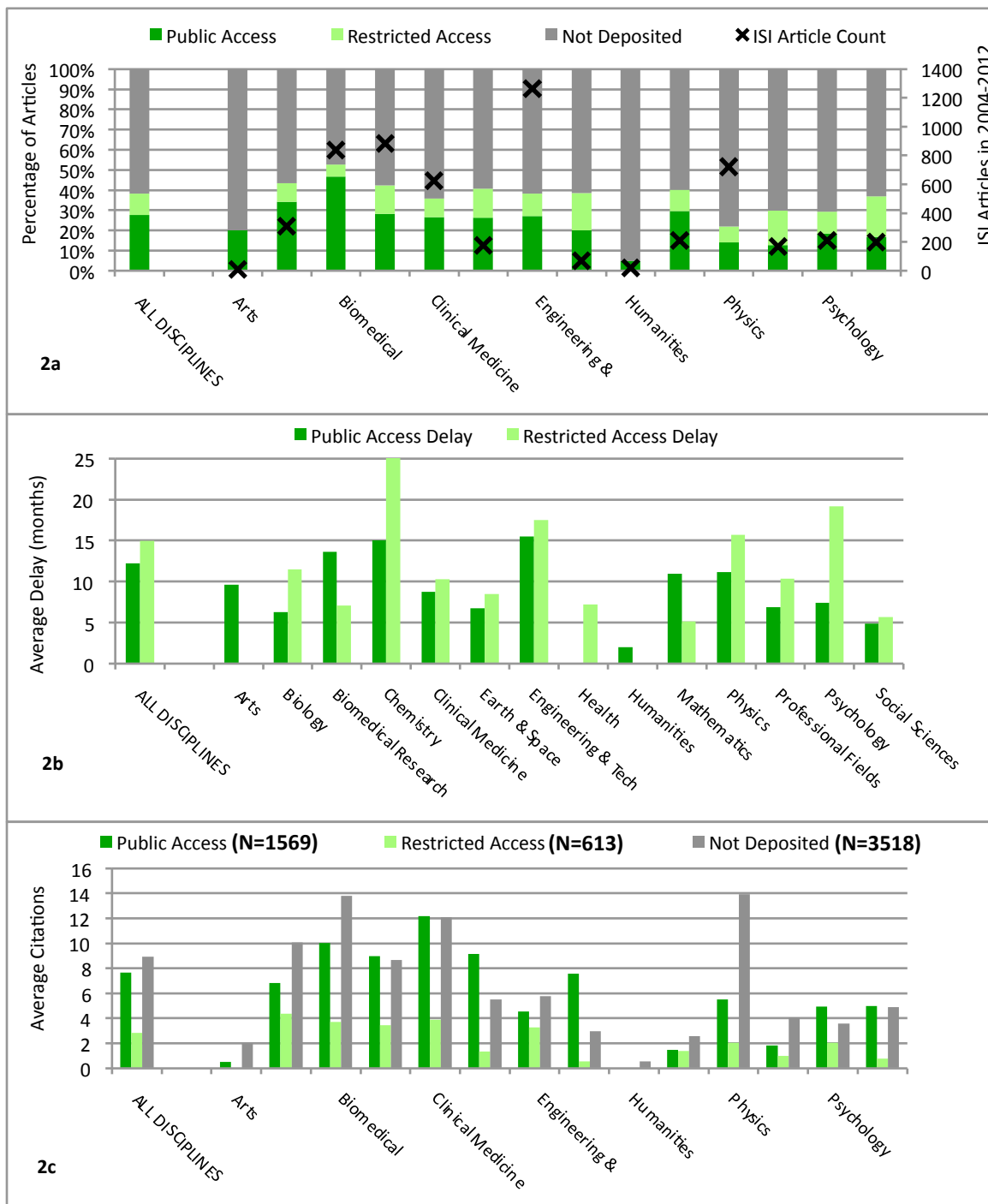


Figure 2. Minho Percentage Deposit, Deposit Delay, and Average Citations, by Discipline.

2a (top): X's indicate number of U Minho ISI-indexed articles published in 2004-2012 (scale on right). Bars are percentage by discipline of articles made Public Access (PA, dark green), Restricted Access (RA, light green) or not deposited (gray) by September 2013.

2b (middle): Average delay (in months) for PA and RA deposits.

2c (bottom): Average citation counts for PA, RA and not-deposited

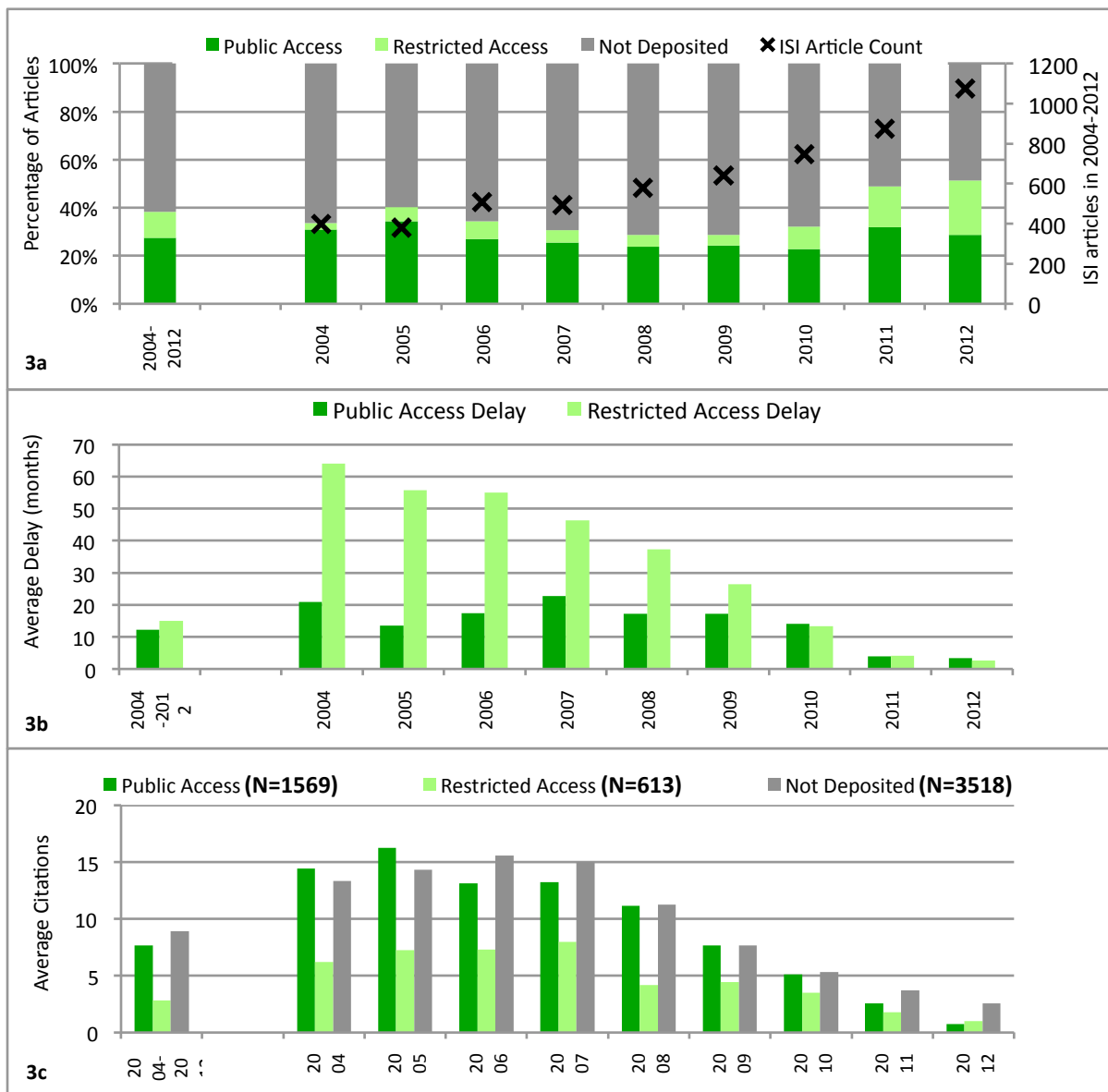


Figure 3. Minho Percentage Deposit, Deposit Delay, and Average Citations, by Year.

3a (top): X's indicate number of U Minho ISI-indexed articles published in 2004-2012 (scale on right). Bars are percentage by publication year of articles made Public Access (PA, dark green), Restricted Access (RA, light green) or not deposited (gray) by September 2013.

3b (middle): Average delay (in months) for PA and RA deposits.

3c (bottom): Average citation counts for PA, RA and not-deposited

3. Deposit Percentage for ISI-indexed Articles. The total number of Minho articles published between 2004 and 2012 and indexed by Thomson-Reuters ISI database was about 5700. The metadata for those articles were extracted from the Minho Institutional Repository (IR) in September 2013. **Figure 2** (& **Table 1**,

Appendix 1) show that of the total ISI sample of 2004-2012 articles, 39% had been deposited as of September 2013, when this analysis was done: 28% as Public Access Full-Texts (PA) and 11% as Restricted Access Full-Texts (RA). (There are no metadata-only records, because the Minho IR only accepts deposits with full-text.) The pattern is similar across disciplines (although Minho's ISI articles in Arts and Humanities are probably too few to conclude that their percentage Green is lowest). The low deposit level in physics is almost certainly due to the fact that physicists (everywhere) have already been depositing, unmandated, in [Arxiv](#) since 1991 and are hence not too keen on having to deposit also in their institutional repositories. (The simple solution is to [import](#) Minho deposits from Arxiv, and there is already repository software to do that.)

Figure 3 (& Table 4, Appendix 2) show percent PA and RA for 2004-2012, by year. The individual years' percentages are similar to the cumulative total for 2004-2012. Percent PA was relatively high in the first 2 years following the mandate (c. 29%), fell (to c. 24% 2007-2010, and then rose again (to c. 30%) for 2011-2012. The rise was probably the beginning of the effect of adopting the stronger Liège model mandate in 2011 (Rentier & Thirion 2011).

4. Comparison with UK. Minho's percentage PA and RA are both much higher than the average for UK IRs (for both the 26 mandated and the 73 non-mandated repositories that were analyzed for comparison) (**Figure 4 & Table 7, Appendix 3**). Minho has almost the same percent PA as U Liège for publication year 2012 (measured in September 2013). However, percent RA is much lower for Minho than for Liège (23% vs 51%). This is something that the upgrade to the stronger Liège mandate in 2011 is probably now in the process of remedying, with publications deposited in RepositóriUM (whether PA or RA) being now the only source of information for institutional and individual research performance assessment. University of Surrey, which has a much weaker mandate than the Liège model generates even fewer deposits, even though it is one of the UK mandated repositories with the highest deposit rate. University of Lancaster, with one of the UK's highest non-mandated deposit rates, is still lower.

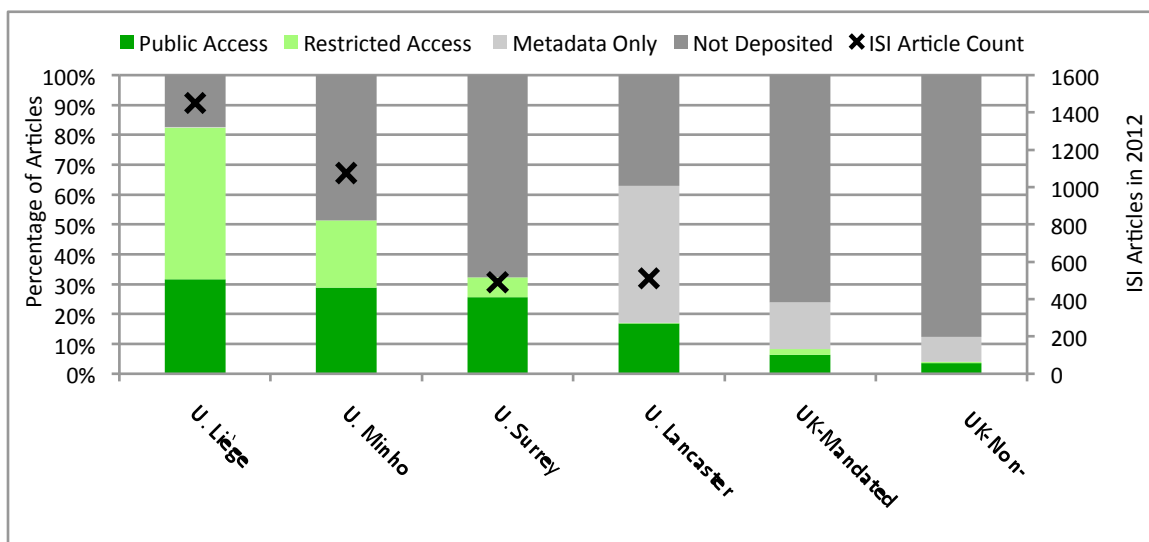


Figure 4. Comparing Deposit Rates. Percent Public Access deposits (PA), Restricted Access deposits (RA), Metadata Only and Not Deposited for the universities of Liège, Minho, Surrey (Mandated) and Lancaster (Non-Mandated) as well as the averages for 26 Mandated and 73 Non-Mandated UK repositories that were analyzed for publication year 2012

5. Importance of Immediate Access. Over 60% of journals endorse immediate, unembargoed Green OA self-archiving today. So there is really no reason why PA deposits should not all be at least 60%. In addition, most of the 40% of publishers who embargo Green OA have now been forced, because of the demand for OA from the author community, to reduce their embargoes to a year or even six months. Yet PA after 6-12 months or more is still not Open Access (OA), which means *immediate access*: It is delayed access (DA). And delaying access matters, because lost research access means lost research uptake, usage, and impact as **Figure 5** shows -- hence lost (not just delayed) research applications, productivity and progress. Physics papers that are not made OA before or at publication never reach the citation level of OA papers; this is not just an OA effect, it is also an early access effect (Gentil-Beccot et al 2010). In a fast-moving field, one must strike while the iron is hot. (Note that even though there is no doubt an element of author self-selection in the Gentil-Beccot et al effect -- with authors of better papers more likely to make them OA -- the effect is too big to be just an author bias: the OA citation advantage is still present, though smaller in size, even for papers that are OA because OA self-archiving was mandatory, rather than just author self-selection; Gargouri et al 2012a.)

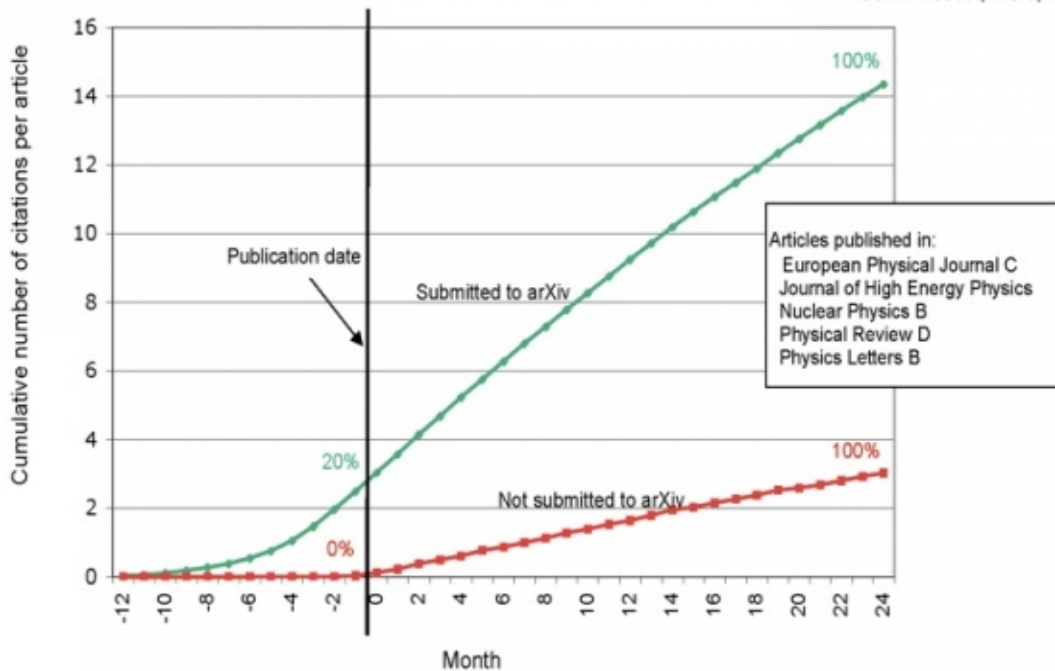


Figure 5. Citation growth for papers made OA earlier (Green) vs. later (red). Note that the red curve never catches up with the green curve. (Figure and data from Gentil-Beccot et al 2010.)

6. The OA Citation Advantage. Gargouri et al (2010) showed (as [many other studies](#) have shown) that OA articles are cited significantly more than non-OA articles. The comparison was done by a robot that trawled the web using Google and Google Scholar to search for OA versions of all ISI-indexed articles. The citation counts for those articles were then compared to matched control non-OA articles in the same journal and year. For the present study it was not possible to replicate that procedure for Minho (because Google now blocks such robot trawling) so **Figure 2c** and **3c** only compares the citation counts for PA, RA and non-deposited articles. The meaningfulness of this comparison is very limited, because there is no matching between OA and non-OA citation counts, but two interesting trends are worth noting: Despite all the uncontrolled confounding factors, Minho's PA articles tend to have higher citation counts than non-OA articles, with the prominent exception of biomedical articles and physics articles. For the physics articles this is almost certainly because most of the non-deposited ones are in fact deposited in Arxiv, hence OA. And a significant portion of the biomedical articles are probably in "Gold" OA journals (Gargouri et al 2012b), which are likewise not deposited, but OA. Other non-deposited Minho

articles may also be OA, elsewhere on the Web. Minho's 2011 upgrade to the Liège model mandate should now be ensuring a much higher percentage deposit, which should in turn be increasingly reflected in a citation advantage even for the raw comparison between PA deposits and non-deposited articles. This citation advantage should also be increased by the earlier depositing induced by Minho's 2011 upgrade to the Liège model.

Because of the importance of early access for research uptake and citation impact, it is important to analyze the *timing* of deposit: PA a year or more after publication is not OA but *delayed access*, DA, and less likely to contribute to research progress and impact, especially in fast-moving fields.

7. Deposit Delay. By subtracting the date articles were published (as indicated by ISI) from the date they were deposited (as indicated by Minho's IR), we calculated the average deposit delay in months. **Figure 2b (& Table 2, Appendix 1)** show that across the entire span of 2004-2012, the average delay for PA deposits was about 12 months (varying from 1 week before publication for Health to 16 months after publication for Engineering & Technology). For RA deposits the average delay was 15 months (from 5 months for Mathematics to 25 months for Chemistry).

Several independent factors contribute to this time course: (1) When a mandate is first adopted, many authors deposit not only their current articles but also their older ones. (2) Similarly, in 2011, when the Minho mandate was upgraded to the Liège model, there was again a spike in both current and older deposits. (3) There is also a spike in current deposits as annual review time approaches (because of the mandate's requirement to deposit in order to be considered for performance review); for tardy authors, the delay between acceptance date and deposit can hence be almost a year for papers published early in the review cycle.

8. Monitoring Timely Compliance. This problem can be remedied in both institutional and funder mandates by a procedure that systematically monitors timely compliance by comparing each date of deposit with the date of the author's acceptance letter. The mandate itself may already be sufficient to ensure timely deposit: Since 2004, Minho deposit delays have been decreasing for both PA deposits (from 21 months for publication year 2004 to 3 months for 2012) and RA deposits (from 64 months for 2004 to 3 months for 2012) (**Figure 2b, 3b, 6 and 7**). For Liège, the average deposit date is already earlier than the date of publication for both PA and RA deposits (**Figure 7**).

The Minho distribution of 2004-2012 articles by deposit delay (**Figure 6**) shows that, there are more PA than RA deposits at each delay level. About 80% of all deposits are made between -5 and +22 months from the date of publication.

9. Comparing Deposit Timing. The distribution of 2012 articles by deposit delay (**Figure 7**) shows that PA and RA deposits are about the same at each delay level. About 90% of all Minho PA deposits (**7a**) are made between -4 and +12 months from the date of publication. In contrast, deposits are made earlier in Liège (**7b**), where about 90% of all PA deposits are made between -8 and +12 months from the date of publication. Minho deposit delays are currently on average longer than those of Surrey (**7c**) (one of the mandated UK institutions with the highest number of PA deposits) where 90% of PA deposits are made between -1 and +12 months from the date of publication. Minho average delays are about the same as those for Lancaster (**7d**) (one of the UK Non-Mandated institutions with the highest number of PA deposits) where 90% of PA deposits are made between -8 and +9 months from the date of publication. For Liège, where deposit is mandatory immediately upon publication -- and for Minho since 2011 -- deposit is the only means of submitting publications for research performance review and research funding. Hence there is every reason to expect that Minho delays are now shrinking to those of Liege.

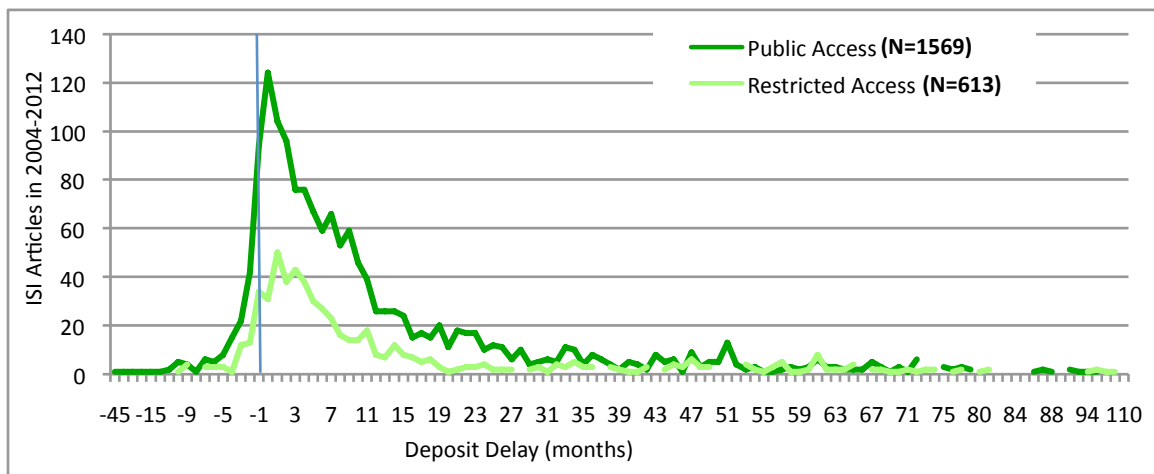


Figure 6. Minho Distribution of 2004-2012 delay (months) for PA and RA deposits. (Some of these occurred as a result of the 2011 increase in deposits with the upgrade to the Liège model mandate, and were accompanied by retroactive deposits of earlier year publications.)

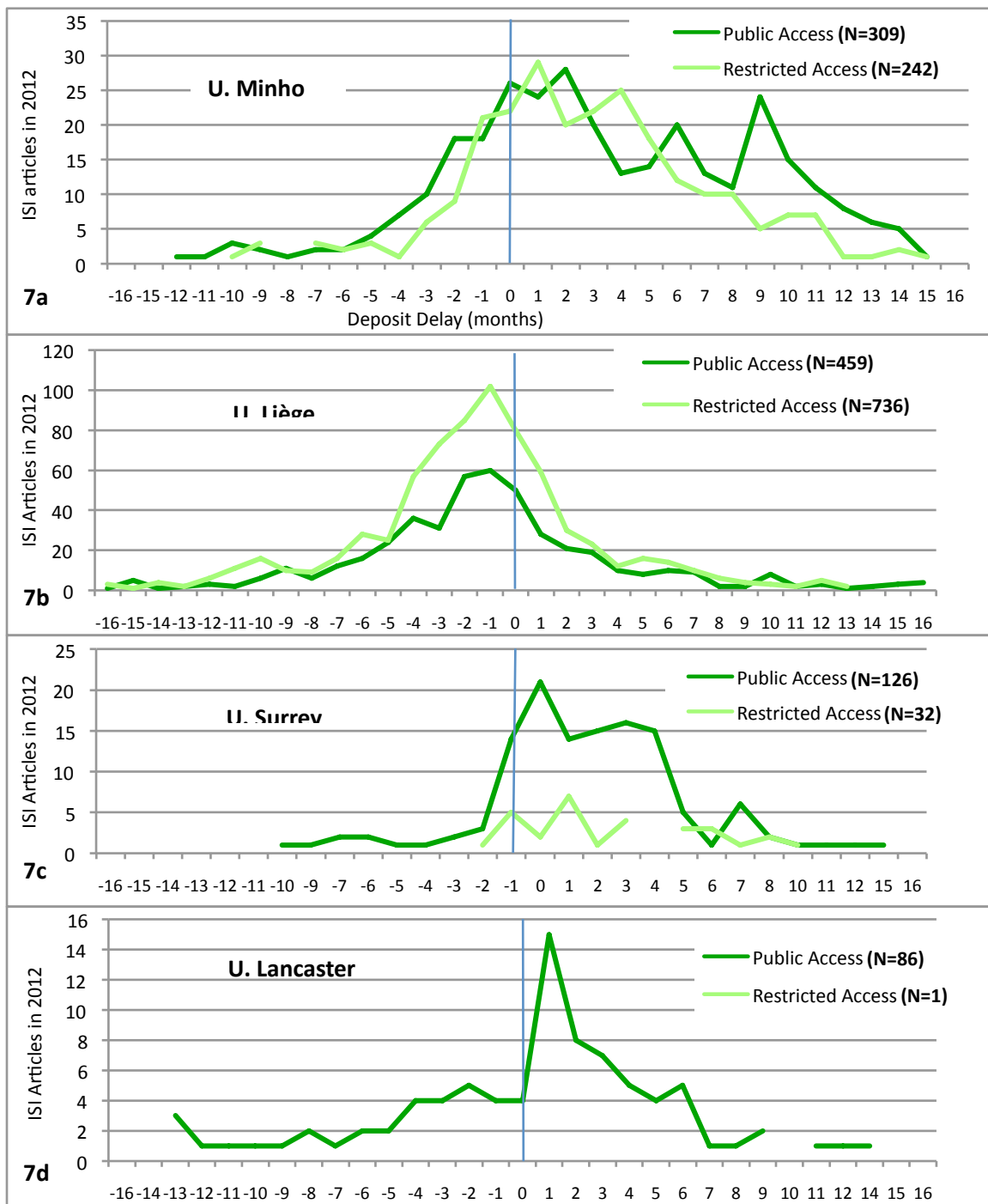


Figure 7. Distribution of 2012 delay (months) for PA and RA deposits for U. Minho, U. Liège, U. Surrey (Mandated) & U. Lancaster (Non-Mandated). Note that Liège's averages (7b) precede the publication date for both PA and RA deposits.

10. Other Kinds of Content. This analysis has been on the OA movement's primary target: refereed journal articles. We close with an analysis of all types of documents deposited in Minho's repository (N= 23738, for publication years 1976-

2013). **Figure 8** shows that articles and conference papers are the most frequent document type (35% each), followed by Master theses (11%), Doctoral theses (5%) and Book Chapters (5%).

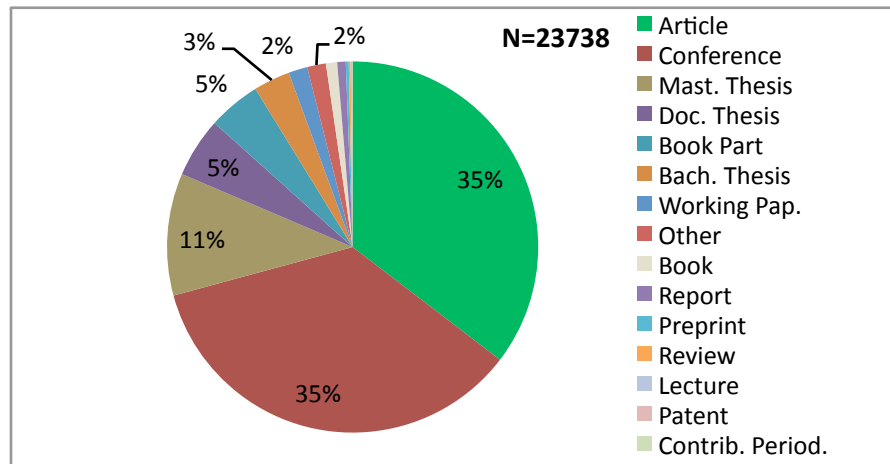


Figure 8. Types of Documents Deposited, 1976-2013

Figure 9 and **Figure 10** show that for each of type of document, Minho deposit rates are rising year by year, since the mandate was adopted in 2004. PA and RA deposits have increased, respectively, from 354 and 22 for publication year 2004 to 809 and 519 for 2012. (Note that the 2013 dip is due to the deposit delay effect.)

Figure 11 shows that the rate of deposit for all kinds of document is increasing each year.

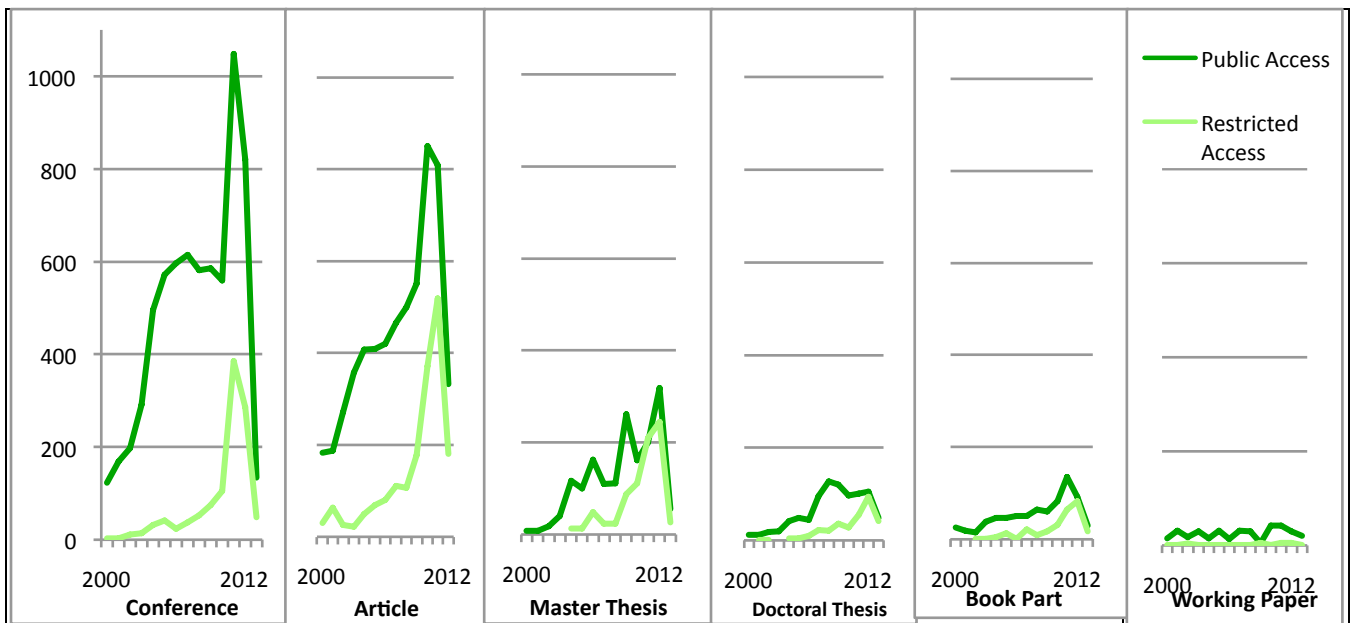


Figure 9. Number of Documents Deposited by Publication Year (2000-2013) and Document Type (only most frequent types of documents are included)

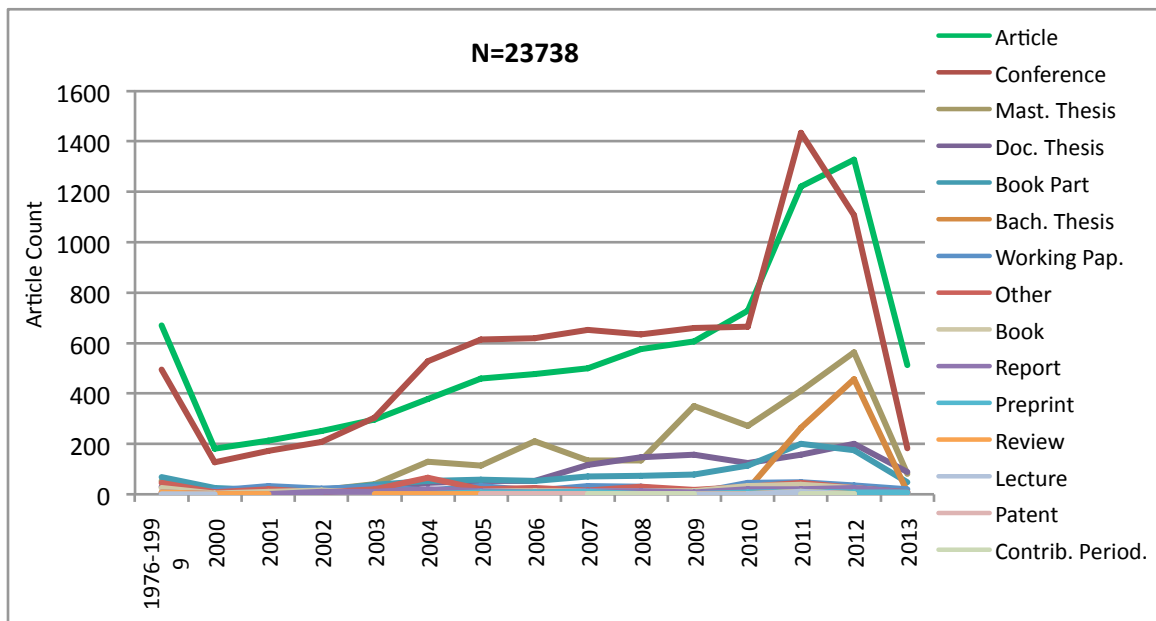


Figure 10. Total Number of Documents Deposited as of September 2013, by Publication Year 1976-2013

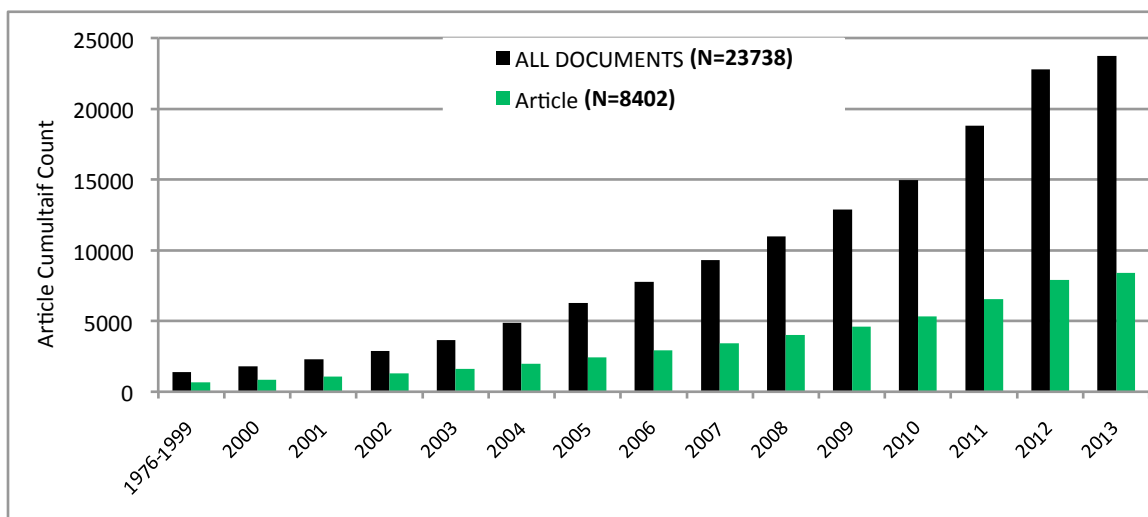


Figure 11. Cumulative Deposits Documents in Minho Repository for articles (blue) and all other kinds of document.

11. Summary. University of Minho's 2004 Green OA mandate has been very successful and with its 2011 upgrade to the Liège model, it has joined the ranks of the small number of mandates worldwide that are optimal (Gargouri et 2012a). It is to be hoped that with the help of complementary, convergent funder mandates (such as [HEFCE](#) in the UK and [FRNS](#) in Belgium), more and more institutions will now adopt the Liège/Minho mandate. Once that is done globally, 100% (Green) OA will not be far behind.

References

Gargouri, Y., Lariviere, V., Gingras, Y., Brody, T., Carr, L., & Harnad, S. (2012a). Testing the Finch Hypothesis on Green OA Mandate Ineffectiveness. *arXiv preprint arXiv:1210.8174*. <http://arxiv.org/pdf/1210.8174>

Gargouri, Yassine, Lariviere, Vincent, Gingras, Yves, Carr, Les and Harnad, Stevan (2012b) Green and Gold Open Access Percentages and Growth, by Discipline. In: *17th International Conference on Science and Technology Indicators (STI)*, 5-8 September, 2012, Montreal, Quebec, Canada, Montréal. <http://eprints.soton.ac.uk/340294/>

Gargouri, Y., Hajjem, C., Lariviere, V., Gingras, Y., Brody, T., Carr, L. and Harnad, S. (2010) Self-Selected or Mandated, Open Access Increases Citation Impact for Higher Quality Research. *PLOS ONE* 5 (10) e13636 <http://eprints.ecs.soton.ac.uk/18493/>

Gentil-Beccot, A., Mele, S., & Brooks, T. C. (2010). Citing and reading behaviours in high-energy physics. *Scientometrics*, 84(2), 345-355. <http://www.slac.stanford.edu/pubs/slacpubs/13500/slac-pub-13693.pdf>

Rentier, B., & Thirion, P. (2011). *The Liège ORBi model: Mandatory policy without rights retention but linked to assessment processes*. <http://orbi.ulq.ac.be/handle/2268/102031>

Sale, A., Couture, M., Rodrigues, E., Carr, L. and Harnad, S. (2012) Open Access Mandates and the "Fair Dealing" Button. In: *Dynamic Fair Dealing: Creating Canadian Culture Online* (Rosemary J. Coombe & Darren Wershler, Eds.) <http://eprints.ecs.soton.ac.uk/18511/>

Appendix 1

Percentage of U Minho ISI-indexed articles published in 2004-2012 deposited in Public Access [PA], Restricted Access [RA] or not deposited: Totals and by Discipline (measured September 2013)

Discipline	Total	Not Deposited		Public Access		Restricted Access	
		N	%	N	%	N	%
ALL DISCIPLINES	5700	3518	62%	1569	28%	613	11%
Arts	10	8	80%	2	20%		0%
Biology	308	174	56%	105	34%	29	9%
Biomedical Research	838	397	47%	390	47%	51	6%
Chemistry	883	509	58%	249	28%	125	14%
Clinical Medicine	624	401	64%	165	26%	58	9%
Earth and Space	175	104	59%	46	26%	25	14%
Engineering & Tech	1264	782	62%	341	27%	141	11%
Health	70	43	61%	14	20%	13	19%
Humanities	21	20	95%	1	5%		0%
Mathematics	209	125	60%	62	30%	22	11%
Physics	722	564	78%	101	14%	57	8%
Professional Fields	167	117	70%	21	13%	29	17%
Psychology	211	149	71%	39	18%	23	11%
Social Sciences	198	125	63%	33	17%	40	20%

Table 2. Average Delay (in months) for Public Access and Restricted Access deposits for publication span 2004-2012 articles by Discipline

Discipline	TOTAL			Public Access			Restricted Access		
	N	Avg Delay	St. Dev.	N	Avg Delay	St. Dev.	N	Avg Delay	St. Dev.
ALL DISCIPLINES	2182	13.0	19.4	1569	12.2	18.2	613	15.0	22.0
Arts	2	9.6	19.5	2	9.6	19.5			
Biology	134	7.4	10.0	105	6.3	8.3	29	11.5	14.2
Biomedical Research	441	12.9	17.6	390	13.6	18.3	51	7.1	8.2
Chemistry	374	18.4	23.9	249	15.0	20.7	125	25.1	28.1
Clinical Medicine	223	9.1	12.8	165	8.7	12.1	58	10.3	14.9
Earth & Space	71	7.3	12.8	46	6.7	9.2	25	8.5	17.7
Engineering & Tech	482	16.0	22.0	341	15.5	21.6	141	17.5	22.9
Health	27	3.4	12.4	14	-0.2	4.1	13	7.2	16.8
Humanities	1	2.0		1	2.0				
Mathematics	84	9.4	16.1	62	11.0	17.8	22	5.1	8.8
Physics	158	12.8	18.4	101	11.2	15.8	57	15.7	22.2
Professional Fields	50	8.9	17.5	21	6.9	13.1	29	10.4	20.1
Psychology	62	11.8	24.1	39	7.4	19.1	23	19.1	29.7
Social Sciences	73	5.3	12.5	33	4.9	13.7	40	5.7	11.5

Table 3. Average Citations for Public Access, Restricted Access deposits and Not Deposited for publication span 2004-2012 articles by Discipline

Discipline	TOTAL N	Avg Cit	Public Access			Restricted Access			Not Deposited		
			N	Avg Cit	St. Dev.	N	Avg Cit	St. Dev.	N	Avg Cit	St. Dev.
ALL DISCIPLINES	5700	7.9	1569	7.7	7.0	613	2.8	13.1	3518	8.9	26.6
Arts	10	1.7	2	0.5				0.7	8	2.0	3.6
Biology	308	8.4	105	6.8	5.4	29	4.4	8.8	174	10.1	17.1
Biomedical Research	838	11.4	390	10.0	6.6	51	3.7	15.9	397	13.8	46.8
Chemistry	883	8.0	249	9.0	5.2	125	3.4	10.9	509	8.7	11.2
Clinical Medicine	624	11.4	165	12.2	8.0	58	3.9	19.6	401	12.1	19.4
Earth and Space	175	5.9	46	9.2	2.1	25	1.3	16.1	104	5.5	7.4
Engineering & Technology	1264	5.2	341	4.5	11.1	141	3.3	7.5	782	5.8	9.6
Health	70	3.3	14	7.6	1.4	13	0.5	18.5	43	3.0	5.0
Humanities	21	0.5	1	0.0					20	0.6	2.0
Mathematics	209	2.1	62	1.5	2.3	22	1.4	2.8	125	2.6	4.1
Physics	722	11.8	101	5.5	3.9	57	2.1	12.8	564	13.9	46.6
Professional Fields	167	3.3	21	1.9	1.8	29	1.0	4.0	117	4.1	8.2
Psychology	211	3.7	39	4.9	2.3	23	2.0	9.1	149	3.6	5.3
Social Sciences	198	4.1	33	5.0	1.1	40	0.8	7.7	125	4.9	10.8

Appendix 2

Table 4. Percentage of U Minho ISI-indexed articles published in 2004-2012 deposited in Public Access [PA], Restricted Access [RA] or not deposited: Totals and by publication year (measured September 2013)

Publication Year	Total	Not Deposited		Public Access		Restricted Access	
		N	%	N	%	N	%
2004-2012	5700	3518	62%	1569	28%	613	11%
2004	398	264	66%	123	31%	11	3%
2005	380	227	60%	131	34%	22	6%
2006	508	333	66%	137	27%	38	7%
2007	494	343	69%	126	26%	25	5%
2008	580	414	71%	138	24%	28	5%
2009	642	458	71%	155	24%	29	5%
2010	747	507	68%	170	23%	70	9%
2011	876	448	51%	280	32%	148	17%
2012	1075	524	49%	309	29%	242	23%

Table 5. Average Delay (in months) for Public Access and Restricted Access deposits for publication span 2004-2012 articles by publication year

Publication Year	TOTAL			Public Access			Restricted Access		
	N	Avg Delay	St. Dev.	N	Avg Delay	St. Dev.	N	Avg Delay	St. Dev.
2004-2012	2182	13.0	19.4	1569	12.2	18.2	613	15.0	22.0
2004	134	24.4	28.2	123	20.9	24.3	11	64.0	38.4
2005	153	19.6	28.1	131	13.5	22.0	22	55.7	33.8
2006	175	25.5	29.2	137	17.3	26.1	38	55.0	19.3
2007	151	26.6	26.2	126	22.7	26.5	25	46.3	12.7
2008	166	20.7	20.3	138	17.3	20.0	28	37.3	12.0
2009	184	18.7	14.6	155	17.2	14.7	29	26.4	11.1
2010	240	13.9	9.7	170	14.1	10.7	70	13.3	6.8
2011	428	4.0	5.3	280	4.0	5.4	148	4.0	5.1
2012	551	3.1	5.1	309	3.4	5.5	242	2.7	4.5

Table 6. Average Citations for Public Access, Restricted Access deposits and Not Deposited for publication span 2004-2012 articles by publication year

Discipline	TOTAL N	Avg Cit	Public Access			Restricted Access			Not Deposited		
			N	Avg Cit	St. Dev.	N	Avg Cit	St. Dev.	N	Avg Cit	St. Dev.
2004-2012	5700	7.9	1569	7.7	7.0	613	2.8	13.1	3518	8.9	26.6
2004	398	13.5	123	14.4	21.2	11	6.2	6.4	264	13.3	17.4
2005	380	14.6	131	16.3	20.7	22	7.2	6.4	227	14.3	19.7
2006	508	14.3	137	13.1	18.7	38	7.3	10.2	333	15.6	49.7
2007	494	14.2	126	13.2	12.5	25	8.0	8.2	343	15.0	31.6
2008	580	10.9	138	11.2	10.5	28	4.2	6.0	414	11.3	44.7
2009	642	7.5	155	7.7	10.5	29	4.4	5.6	458	7.7	12.2
2010	747	5.1	170	5.1	4.5	70	3.5	4.0	507	5.3	9.4
2011	876	3.0	280	2.6	3.6	148	1.8	2.6	448	3.7	8.4
2012	1075	1.7	309	0.7	3.2	242	1.0	8.2	524	2.6	12.0

Appendix 3

Table 7. Percent Public Access deposits (PA), Restricted Access deposits (RA), Metadata Only and Not Deposited for the universities of Liège, Minho, Surrey (Mandated) and Lancaster (Non-Mandated) as well as the averages for 26 Mandated and 73 Non-Mandated UK repositories that were analyzed for publication year 2012

Institution	N	Public Access	Restricted Access	Metadata Only	Not Deposited
U. Liège	1452	31.6%	50.8%	0.2%	17.4%
U. Minho	1075	28.7%	22.5%	0.0%	48.7%
U. Surrey	490	25.7%	6.5%	0.0%	67.8%
U. Lancaster	511	16.8%	0.2%	46.0%	37.0%
UK-Mandated	11995	6.4%	1.8%	15.7%	76.0%
UK-Non-Mandated	24868	3.5%	0.5%	8.3%	87.7%