

Why can't I reuse these tables and figures?

Martin Fenner, Gobbledygook

September 30, 2010

Tables and figures contain the data of a scientific paper in condensed (and often visually appealing) form. This is why they are among the first thing we look at, and why they are often reused when we discuss the paper in a presentation or blog post.

From doi:10.1371/journal.ppat.1001042. Image credit: Thomas J. Hannan, Washington University.

Electronic publication has dramatically simplified the reuse of tables and figures, and therefore reuse has become very common – you probably find reused material in most presentations given in academic institutions or at conferences.

Most authors will probably be happy that their results are disseminated, and reuse is likely to lead to more people reading the full paper and citing the work.

From doi:10.1186/1471-2407-10-286.

But this reuse has two problems. The first problem is copyright. Many journals own the copyright of the papers they publish and don't allow reuse without prior permission. Unfortunately copyright is a complicated issue, and also differs between countries. Most researchers assume "Fair use" when they reuse material, but this might not apply to all situations, e.g. presentations at conferences. And many researchers don't understand that they have often given away the copyright to their own works, so that they can't show a figure from one of their papers without permission.

Many publishers have automated the process of obtaining permissions for copyrighted work, e.g. using the Rightslink system of the Copyright Clearance Center. But it still requires a considerable investment in time (and often money) to obtain all permissions, especially since these are usually one-time permissions only. This combination of unawareness of the details of copyright law and the required extra work means that many researchers probably don't obtain permissions prior to reuse.

The solution to the copyright problems is obviously to use material with a Creative Commons license whenever possible, as I have done in this blog post.

And most Open Access papers are published under this license, so there is plenty of material to choose from.

But there is also a second problem with reusing tables and figures. They were designed to be part of a paper and often look terrible in a presentation, particularly tables.

ID	Type	Measure	Source	Network parameters	PC1	PC2	ρ
1	Citation	Scimago Journal Rank	Scimago/Scopus		-0.974	-0.296	0.556 [*]
2	Citation	Immediacy Index	JCR 2007		1.659	-7.046	0.508 [*]
3	Citation	Closeness Centrality	JCR 2007	Undirected, weighted	0.339	-6.394	0.565 [*]
4	Citation	Cites per doc	Scimago/Scopus		-1.371	-6.192	0.588 [*]
5	Citation	Journal Impact Factor	JCR 2007		-1.854	-5.937	0.582 [*]
6	Citation	Closeness centrality	JCR 2007	Undirected, unweighted	-1.388	-4.827	0.619
7	Citation	Out-degree centrality	JCR 2007	Directed, weighted	-3.191	-4.215	0.642
8	Citation	Out-degree centrality	JCR 2007	Directed, unweighted	-2.793	-4.015	0.640
9	Citation	Degree Centrality	JCR 2007	Undirected, weighted	-4.850	-2.834	0.690
10	Citation	Degree Centrality	JCR 2007	Undirected, unweighted	-4.398	-2.643	0.681
11	Citation	H-index	Scimago/Scopus		-3.326	-2.003	0.681
12	Citation	Scimago Total cites	Scimago/Scopus		-4.906	-1.722	0.712
13	Citation	Journal Cite Probability	JCR 2007		-5.399	-1.647	0.710
14	Citation	In-degree centrality	JCR 2007	Directed, unweighted	-5.392	-1.429	0.717
15	Citation	In-degree centrality	JCR 2007	Directed, weighted	-5.380	-1.554	0.712
16	Citation	PageRank	JCR 2007	Directed, unweighted	-4.476	0.108	0.693
17	Citation	PageRank	JCR 2007	Undirected, unweighted	-4.509	0.731	0.726
18	Citation	PageRank	JCR 2007	Undirected, weighted	-4.190	0.864	0.696
19	Citation	PageRank	JCR 2007	Directed, weighted	-5.333	0.333	0.699
20	Citation	Y-factor	JCR 2007	Directed, weighted	-2.971	0.317	0.657
21	Citation	Betweenness centrality	JCR 2007	Undirected, weighted	-0.462	0.872	0.643
22	Citation	Betweenness centrality	JCR 2007	Undirected, unweighted	-0.474	1.609	0.642
23	Citation	Citation half-life	JCR 2007	/	/	/	0.657
24	Usage	Closeness centrality	MESUR 2007	Undirected, weighted	3.130	2.683	0.793
25	Usage	Closeness centrality	MESUR 2007	Undirected, unweighted	3.100	3.899	0.731
26	Usage	Degree centrality	MESUR 2007	Undirected, unweighted	3.271	3.873	0.729
27	Usage	PageRank	MESUR 2007	Undirected, unweighted	3.327	4.192	0.728
28	Usage	PageRank	MESUR 2007	Directed, unweighted	3.463	4.336	0.727
29	Usage	In-degree centrality	MESUR 2007	Directed, unweighted	3.463	4.015	0.728
30	Usage	Out-degree centrality	MESUR 2007	Directed, unweighted	3.484	3.994	0.727
31	Usage	PageRank	MESUR 2007	Directed, weighted	3.780	4.217	0.710
32	Usage	PageRank	MESUR 2007	Undirected, weighted	3.813	4.223	0.710
33	Usage	Betweenness centrality	MESUR 2007	Undirected, unweighted	3.988	4.271	0.699
34	Usage	Betweenness centrality	MESUR 2007	Undirected, weighted	3.957	3.698	0.693
35	Usage	Degree centrality	MESUR 2007	Undirected, weighted	5.293	3.528	0.683
36	Usage	Out-degree centrality	MESUR 2007	Directed, weighted	5.302	3.518	0.683
37	Usage	In-degree centrality	MESUR 2007	Directed, weighted	5.286	3.531	0.683
38	Usage	Journal Use Probability	MESUR 2007		8.914	1.833	0.593
39	Usage	Usage Impact Factor	MESUR 2007	/	/	/	0.279

Average Spearman rank-order correlations to all other measures are listed under ρ (five lowest values indicated by *).
doi:10.1371/journal.pone.0006022.t001

Figure 1:

From doi:10.1371/journal.pone.0006022.

The solution to this problem is to provide the data behind the table or figure, so that the information can be displayed in a way that makes sense in a presentation. Here we usually have to reduce the amount of information, but it could also mean that we remix the content with other sources. The Creative Commons licenses discussed above are not appropriate for data. Whenever possible, scientific data should be placed in the public domain.

It is important to distinguish the publication of table and figure data from the publication of the whole research dataset. The open questions with the latter (e.g. standard data formats, appropriate repositories, archiving) don't apply to the former. This means that publishers could start providing these data immediately. I'm confident that they would see an increase in paper downloads and citations.

But more importantly I hope this would lead to better presentations in seminars and at conferences.