

Comparing citation database coverage: University of Jyväskylä research publications in Dimensions, Scopus, Web of Science and PubMed

We compare the coverage of research outputs in four citation databases – Dimensions, Scopus, Web of Science and PubMed – for a set of research outputs in a hand-curated Current Research Information System (CRIS) at University of Jyväskylä. The set of research outputs investigated spans years 2007-2019, and includes diverse document types, from peer-reviewed original articles to popularized articles in newspapers.

The query and search procedure is presented in Figure 1. The overall result is described in Table 1.

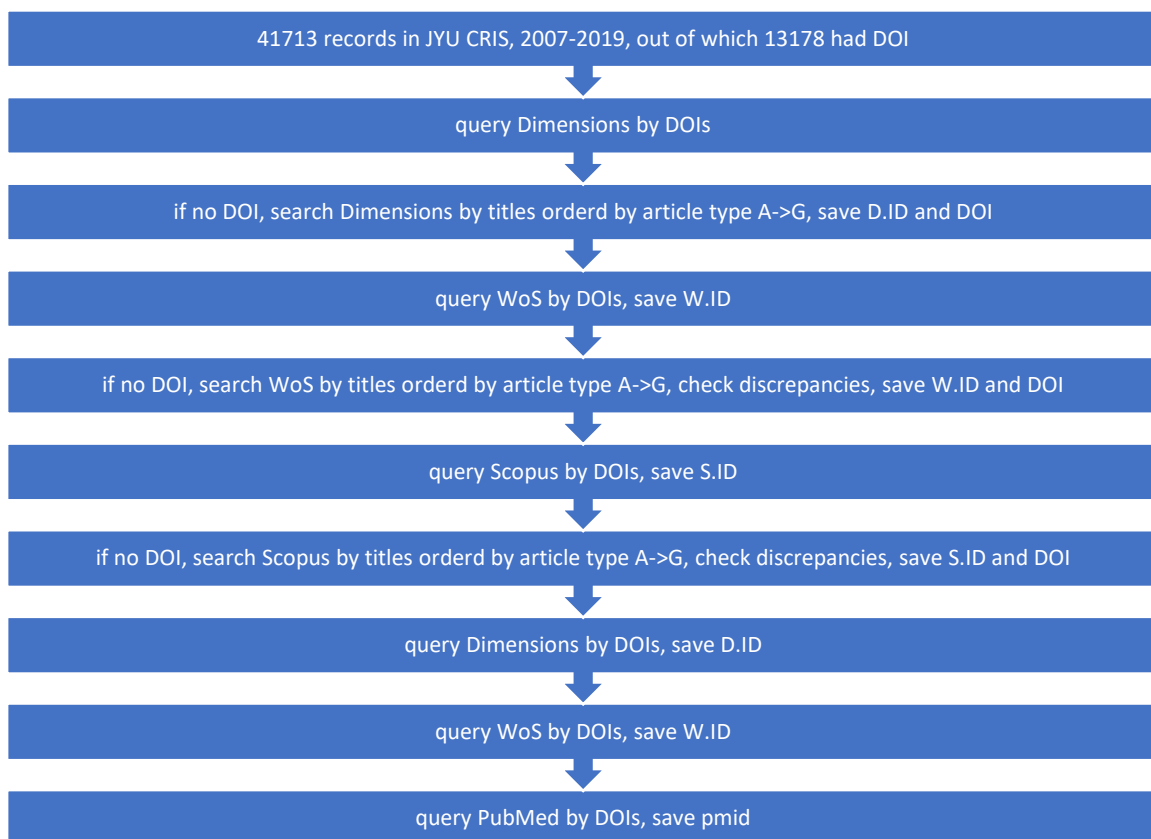


Figure 1.

Table 1. Output sets in JYU CRIS and discovered match counts in Dimensions (D), Scopus (S), Web of Science (W) and PubMed (P)

	JYU CRIS	has DOI				no DOI			
Peer-reviewed scientific articles (A1-A2)	16216	13213 (81%)				3003 (18%)			
		D: >99% (13148)	S: 94% (12422)	W: 89% (11739)	P: 36% (4754)	D: 2% (67)	S: 8% (234)	W: 6% (167)	P: NA
Peer-reviewed chapter in book (A3)	4743	1028 (22%)				3715 (78%)			
		D: 89% (914)	S: 42% (431)	W: 1% (9)	P: <1% (5)	D: 0	S: 2% (72)	W: 0	P: NA
Peer-reviewed conference papers (A4)	3025	1421 (47%)				1604 (53%)			
		D: 99% (1403)	S: 88% (1253)	W: 3% (36)	P: <1% (7)	D: 1% (12)	S: 21% (335)	W: <1% (2)	P: NA
Non-refereed scientific articles (B)	5898	630 (11%)				5234 (89%)			
		D: 96% (604)	S: 51% (323)	W: 40% (254)	P: 15% (96)	D: <1% (4)	S: 1% (66)	W: <1% (5)	P: NA
Scientific books (C)	1391	150 (11%)				1241 (89%)			
		D: 88% (132)	S: 31% (46)	W: 12% (18)	P: 0	D: 0	S: 1% (10)	W: 1% (7)	P: NA
Professional publications (D)	5358	54 (1%)				5304 (99%)			
		D: 87% (47)	S: 35% (19)	W: 4% (2)	P: 0	D: <1% (10)	S: <1% (9)	W: 0	P: NA
Popularized publications (E)	3245	4 (<1%)				3235 (>99%)			
		D: 2	S: 3	W: 0	P: 1	D: 0	S: 2	W: 0	P: NA
Artistic publications (F)	39	0				39 (100%)			
		NA	NA	NA	NA	D: 0	S: 0	W: 0	P: NA
Theses (G)	1798	7 (<1%)				1791 (>99%)			
		D: 6	S: 0	W: 1	P: 0	D: 0	S: 2	W: 4	P: NA
TOTAL	41713	16507 (40%)				25206 (60%)			
		D: 98% (16256)	S: 88% (14497)	W: 73% (12059)	P: 29% (4861)	D: <1% (93)	S: 3% (730)	W: 1% (185)	P: NA

Title search from Web of Science yielded 36 DOI's that did not match with CRIS DOIs. These were manually checked, and errors flagged. Some were due to valid duplicate DOI registrations.

Scopus does not always respect the uniqueness of DOI: for edited volumes, Scopus may have separate entries for each chapter and these can be individually discovered with a title search, but sometimes they all list the same DOI (of the parent volume), both in the web interface and in the API. Sometimes book chapters may have their own unique DOIs.

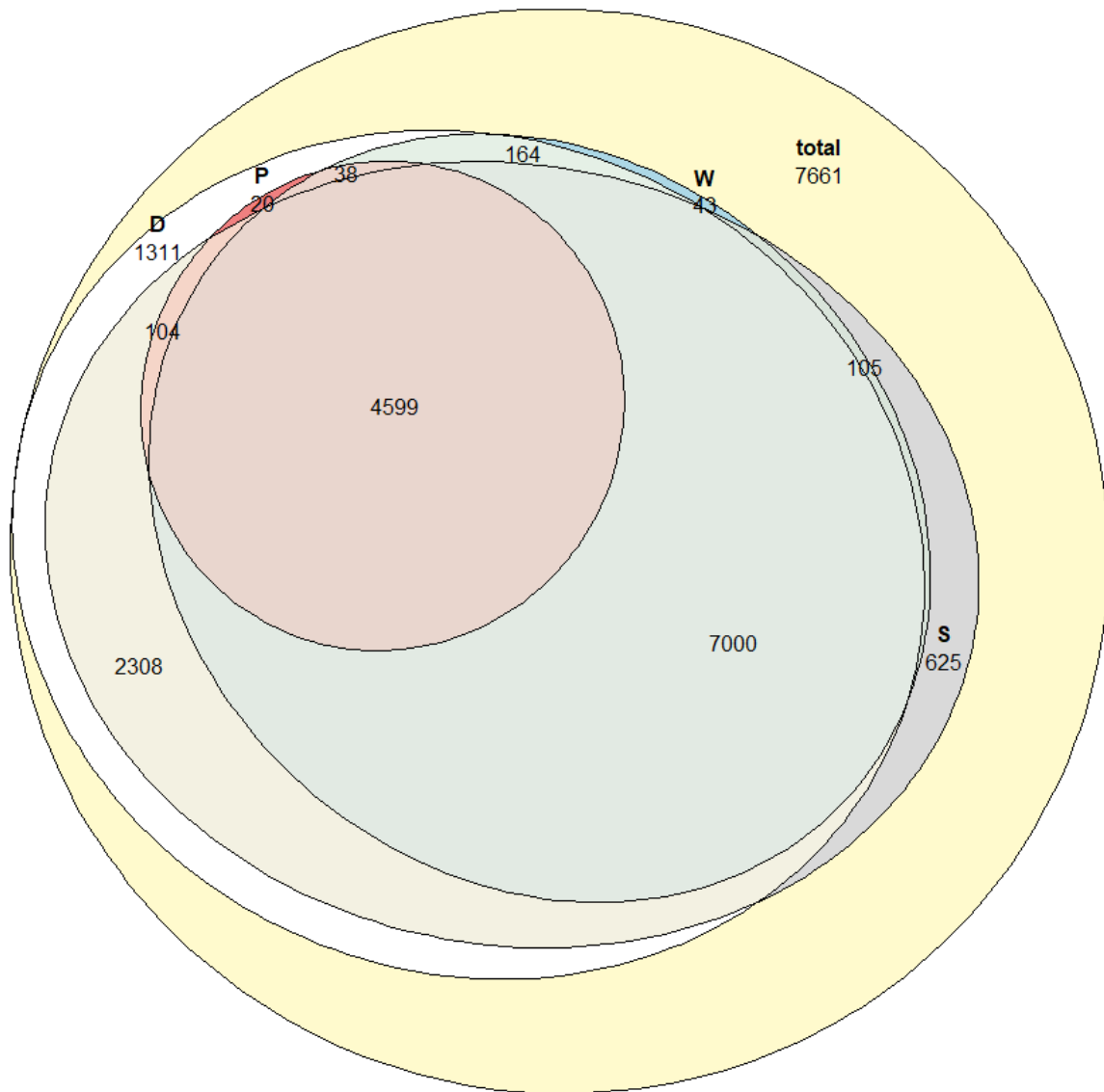


Figure 1a. All peer reviewed outputs: Area-proportional Euler diagram of peer-reviewed output incidence in JYU CRIS (total), Dimensions (D), Scopus (S), Web of Science (W) and PubMed (P). The total ellipse encompasses 23984 outputs.



Figure 1b. All peer reviewed outputs with DOI: Area-proportional Euler diagram of peer-reviewed output incidence in JYU CRIS (total), Dimensions (D), Scopus (S), Web of Science (W) and PubMed (P). The total ellipse encompasses 15662 outputs.

Note that in this set Dimensions almost entirely overlaps with all others, and consequently the diagram algorithm fails to fit accurately. Notably, 133 outputs that can be found in Scopus but not in Dimensions fail to appear in this diagram. This Euler diagram is nonetheless included for completeness and comparison, as it is otherwise accurate. Inclusive output incidence is shown in Fig 1c.

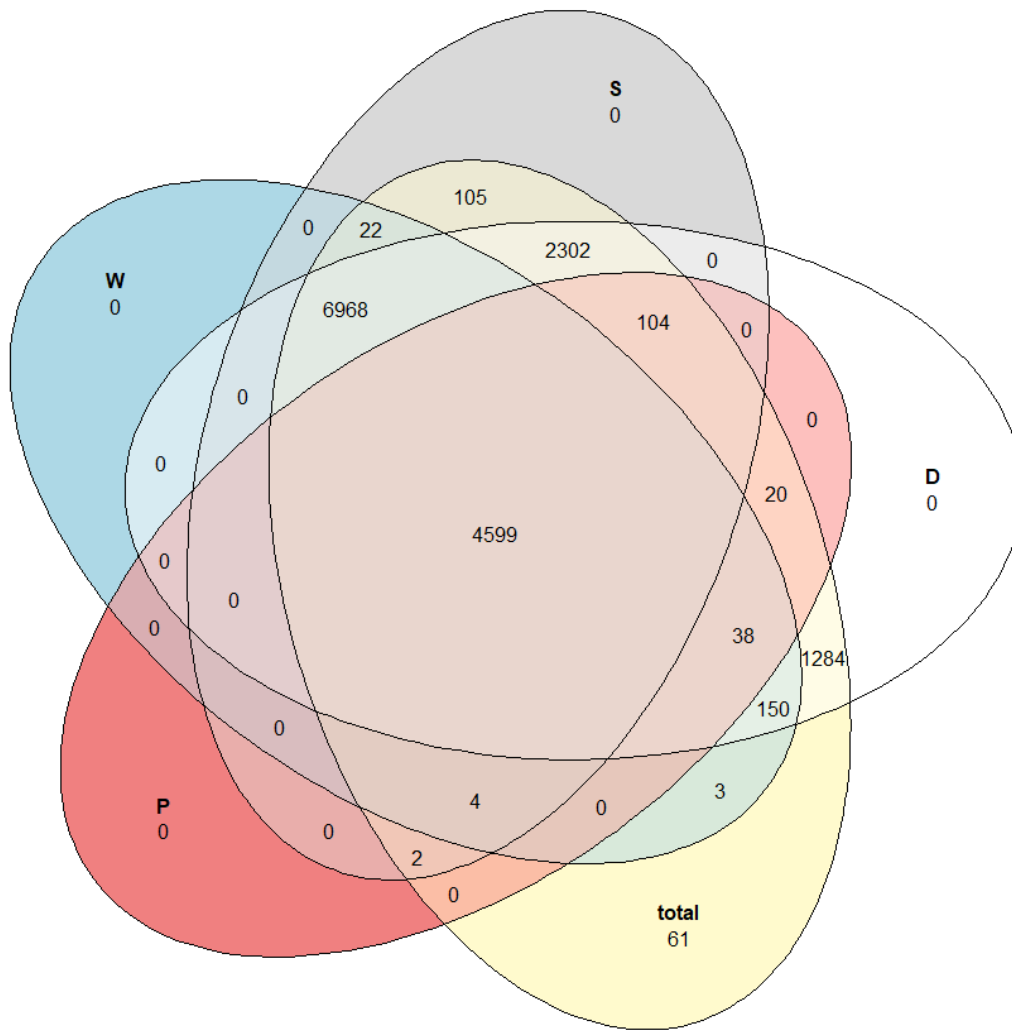


Figure 1c. All peer reviewed outputs with DOI: Venn diagram of peer-reviewed output incidence in JYU CRIS (total), Dimensions (D), Scopus (S), Web of Science (W) and PubMed (P). The total ellipse encompasses 15662 outputs.

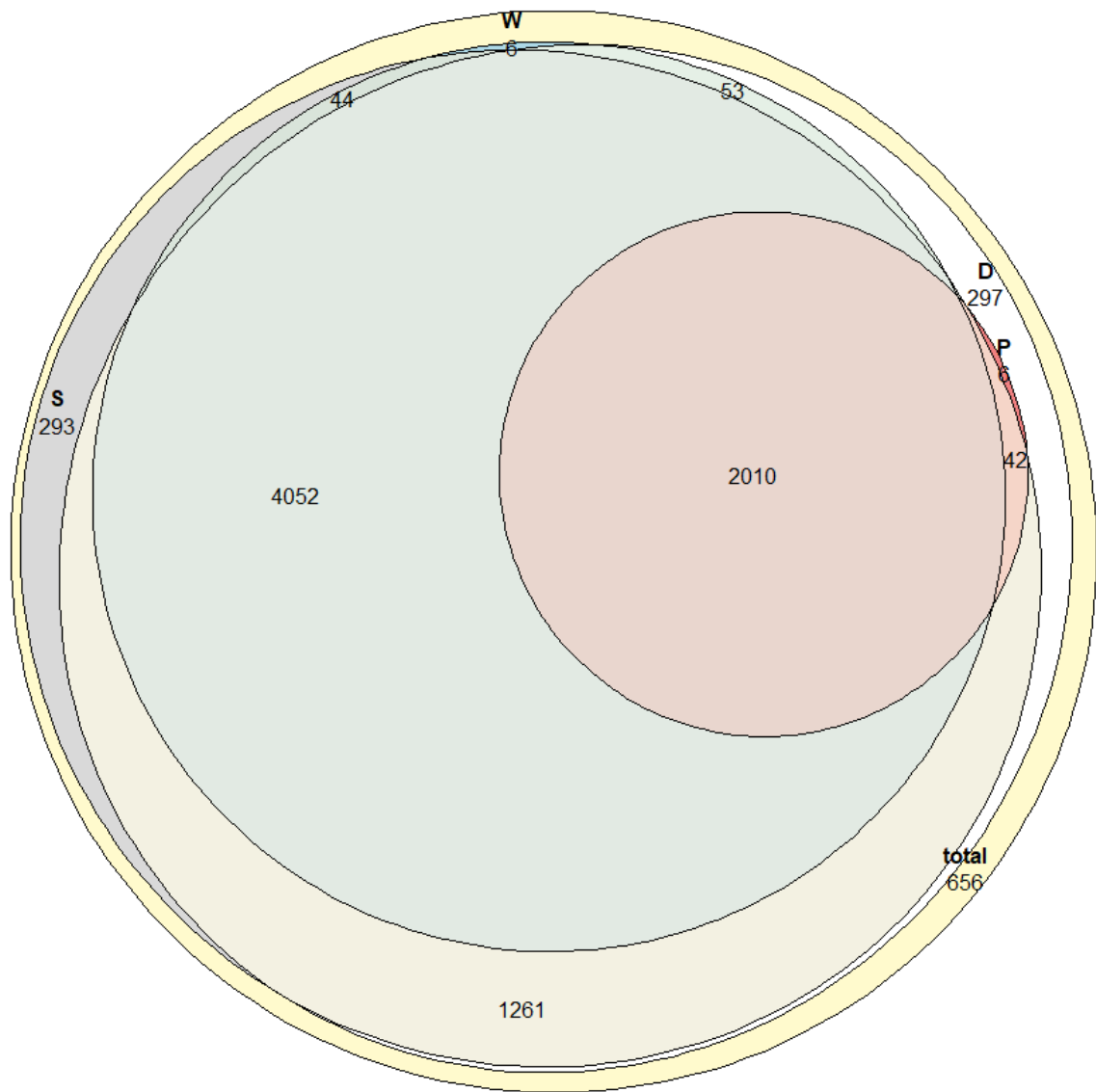


Figure 2. Natural Sciences: Area-proportional Euler diagram of peer-reviewed output incidence in JYU CRIS (total), Dimensions (D), Scopus (S), Web of Science (W) and PubMed (P). The total ellipse encompasses 8735 outputs.

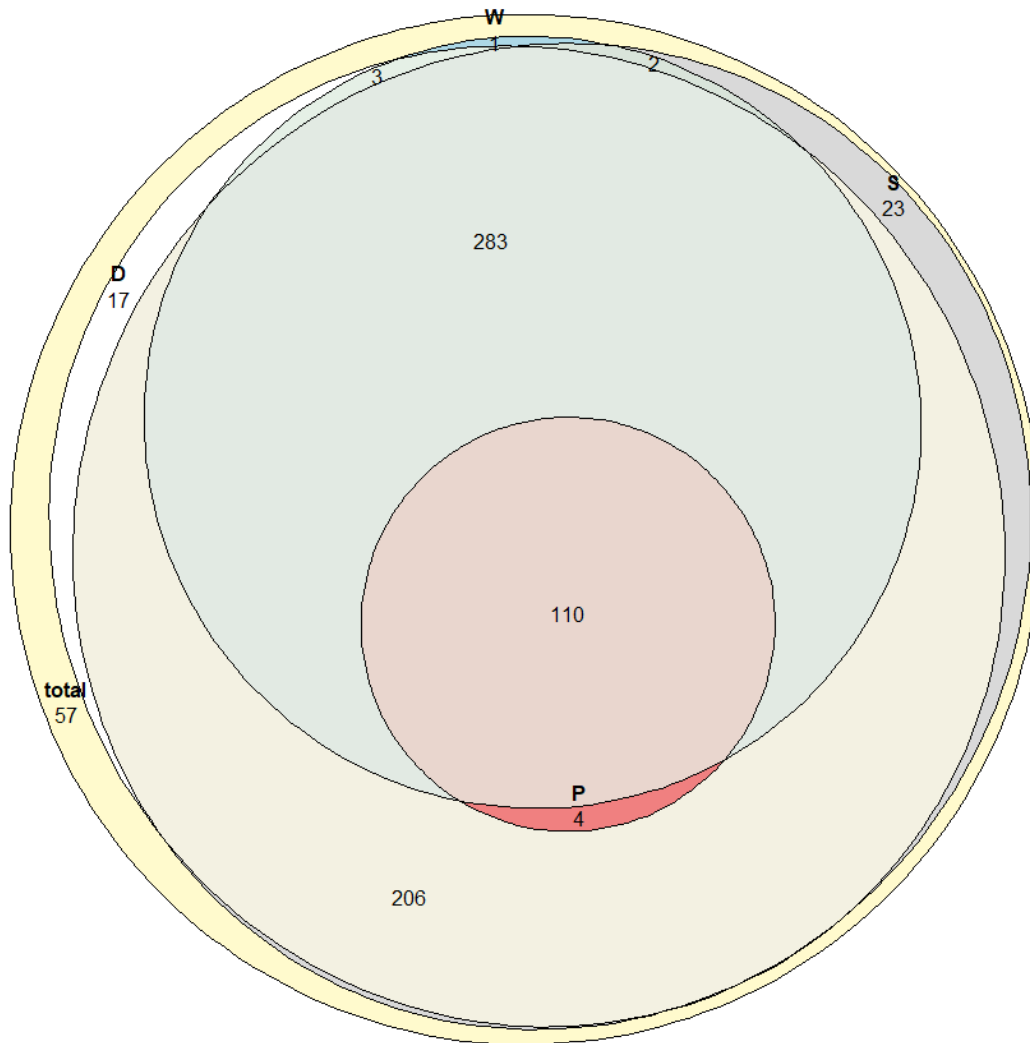


Figure 3. Engineering and Technology: Area-proportional Euler diagram of peer-reviewed output incidence in JYU CRIS (total), Dimensions (D), Scopus (S), Web of Science (W) and PubMed (P). The total ellipse encompasses 706 outputs.



Figure 3. Medical and Health Sciences: Area-proportional Euler diagram of peer-reviewed output incidence in JYU CRIS (total), Dimensions (D), Scopus (S), Web of Science (W) and PubMed (P). The total ellipse encompasses 2646 outputs.

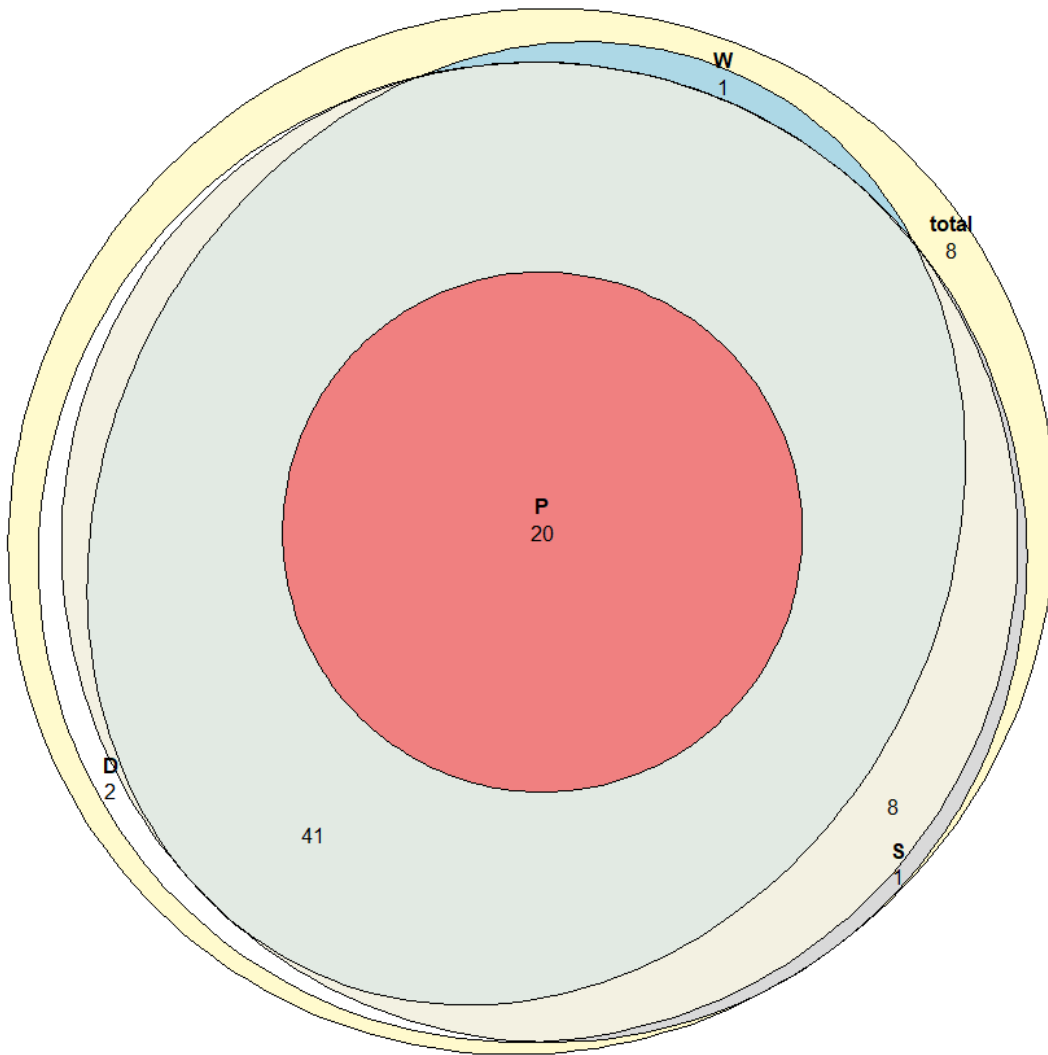


Figure 4. Agricultural Sciences: Area-proportional Euler diagram of peer-reviewed output incidence in JYU CRIS (total), Dimensions (D), Scopus (S), Web of Science (W) and PubMed (P). The total ellipse encompasses 81 outputs.

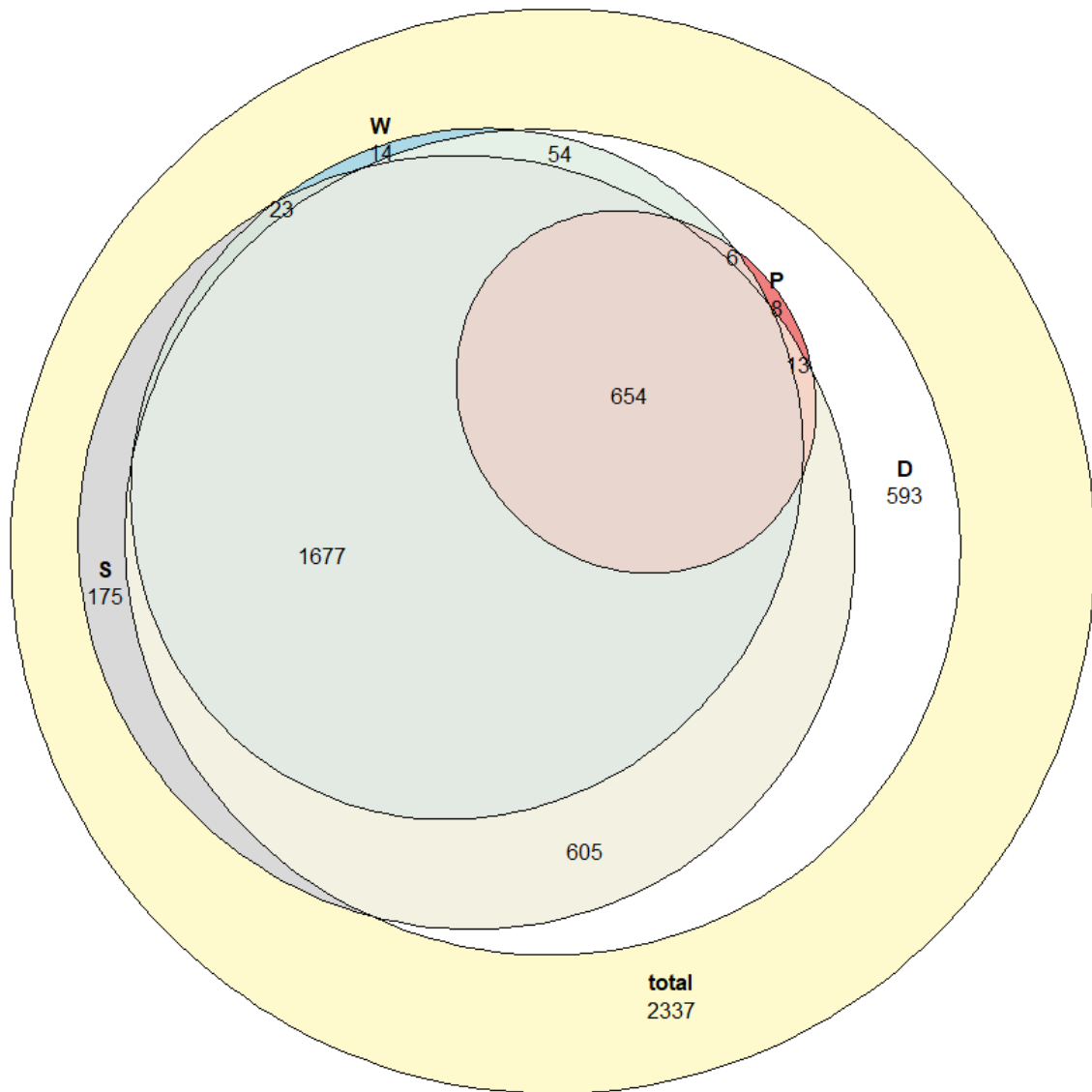


Figure 5. Social Sciences: Area-proportional Euler diagram of peer-reviewed output incidence in JYU CRIS (total), Dimensions (D), Scopus (S), Web of Science (W) and PubMed (P). The total ellipse encompasses 6159 outputs.

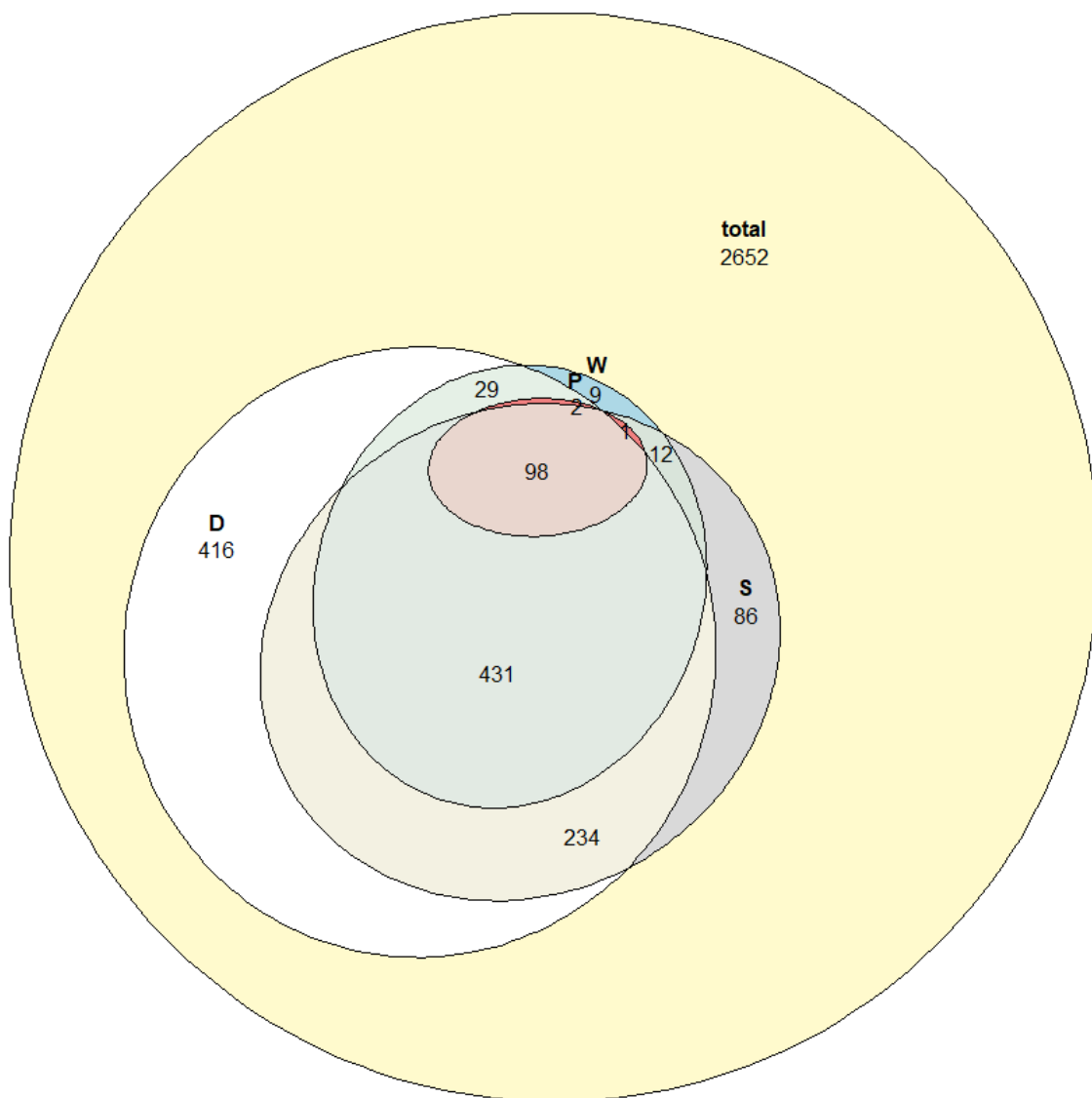


Figure 6. Humanities: Area-proportional Euler diagram of peer-reviewed output incidence in JYU CRIS (total), Dimensions (D), Scopus (S), Web of Science (W) and PubMed (P). The total ellipse encompasses 3970 outputs.

References

Larsson J (2020). `_eulerr`: Area-Proportional Euler and Venn Diagrams with Ellipses_. R package version 6.1.0, <URL: <https://cran.r-project.org/package=eulerr>>.