Music works in the GND and other authority datasets

Laurent Pugin

GNDCon 2.0 – 7 June 2021
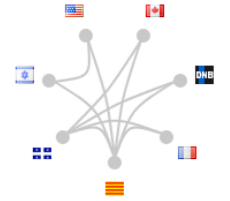
Werknormdaten der Musik

Bibliotheken und Wissenschaft im Gespräch

RISM | DIGITAL CENTER

Datasets

- DNB (started in 2012) - available in RDF
  - 217,055 music works (410,736 works; 5,380 expressions)

- BnF (started in 2011) - available in RDF
  - 172,127 music works

240 _ _ ‡a_Bach ‡b_Johann Sebastian ‡f_1685-1750 ‡t_Musique de chambre ‡s_BWV 1007-1040
100 1 _ ‡a_Bach, Johann Sebastian ‡d_1685-1750 ‡t_Kammermusik
100 1 _ ‡a_Bach, Johann Sebastian, ‡d_1685-1750 ‡t_Musique instrumentale. ‡k_Extraits
100 1 _ ‡a_Bach, Johann Sebastian, ‡d_1685-1750. ‡t_Instrumental music. ‡k_Selections
100 1 _ ‡a_Bach, Johann Sebastian, ‡d_1685-1750. ‡t_Instrumental music. ‡k_Selections
100 1 _ ‡a_Bach, Johann Sebastian, ‡d_1685-1750. ‡t_Musique instrumentale. ‡k_Extraits
100 1 _ ‡a_Bach, Johann Sebastian, ‡d_1685-1750. ‡t_Música instrumental. ‡k_Seleccions

- VIAF
  - 115,570,369 links
  - 855,502 links pointing to or from the DNB of the BnF works

- MusicBrainz (started in 2000)
  - 1,309,201 works
  - MB-1920 with composers born before 1920
    - 236,640 music works
  - MB-Series with curated catalogues
    - 24,198 music works

```
LC|n  82150575
J9U|987007594707705171
BNC@.a12214863
DNB@http://d-nb.info/gnd/1096808099
LC@n82150575
BNF@http://catalogue.bnf.fr/ark:/12148/cb14038626m
DNB|1096808099
B2Q|0000448640
CAOONL|ncf10122328
BNC|a12214863
BNF|14038626
```

- IMSLP (started in 2006)
  - 163,668 music works
  - API (basic data) and HTTP requests

RISM | DIGITAL CENTER

Main questions

✤ Can we compare what these datasets currently contain?

✤ Can we evaluate what their level of linkage currently is?

Laurent Pugin, Claudio Bacciagaluppi:
An Analysis of Musical Work Datasets and their
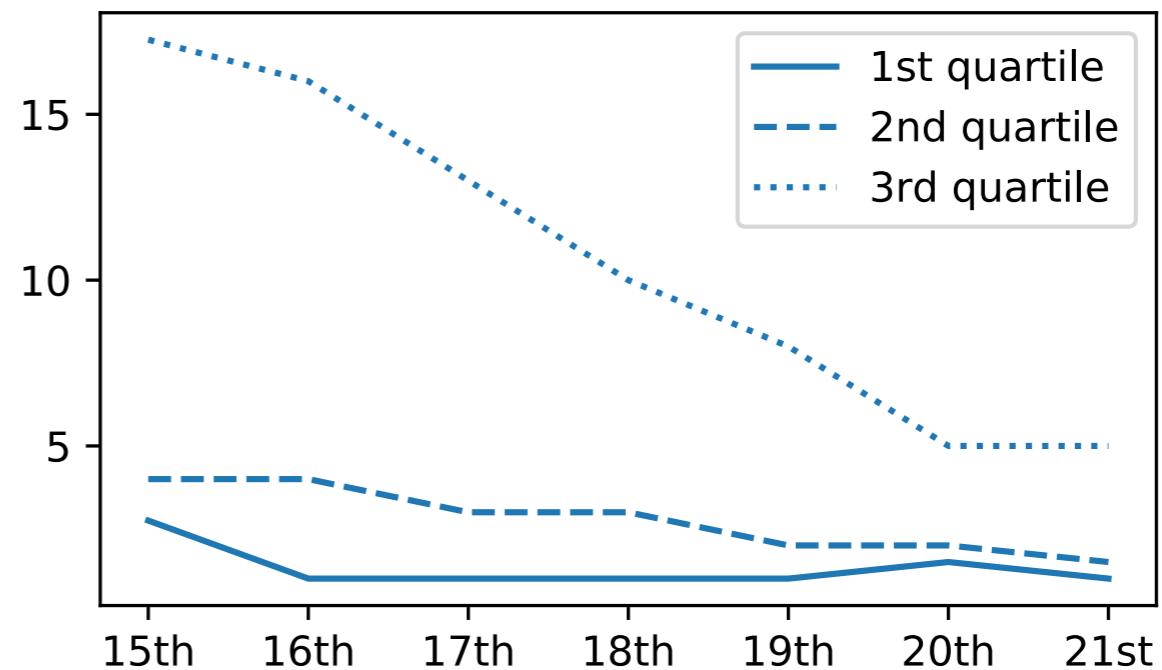Current Level of Linkage. DLfM 2020: 32-39

Analysis by composer

✤ DNB gnd:firstComposer / BnF dcterms:creator

	Works analysed	Nb of composers	Number of works by composer				
			Max	AVG	Q3	Q2	Q1
DNB	208,392	27,878	2,133	7	4	1	1
BnF	173,515	23,332	1,194	7	4	2	1
MB-1920	236,640	10,006	6,501	24	9	3	1
IMSLP	163,668	18,748	2,006	9	6	2	1
MB-Series	24,198	404	1,474	60	50	21	6

Number of works by composer in the datasets with the maximum and average values and the 3rd, 2nd and 1st quartiles (Q3, Q2 and Q1)
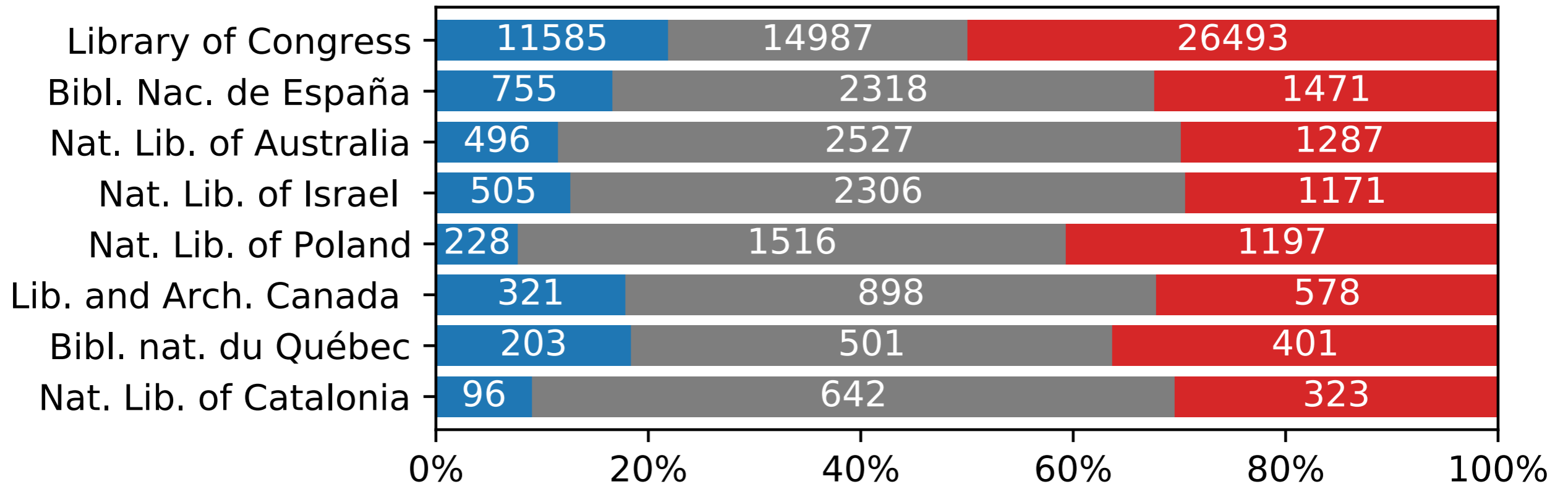
Number of works by composer century by century for the BnF dataset, with 1st,

RISM DIGITAL CENTER

Links between DNB and BnF in VIAF

✤ 18,702 VIAF clusters (11% of BnF and 8.5% of DNB)

Works mapped with other libraries through VIAF

	BNF only	BNF and DNB	DNB only

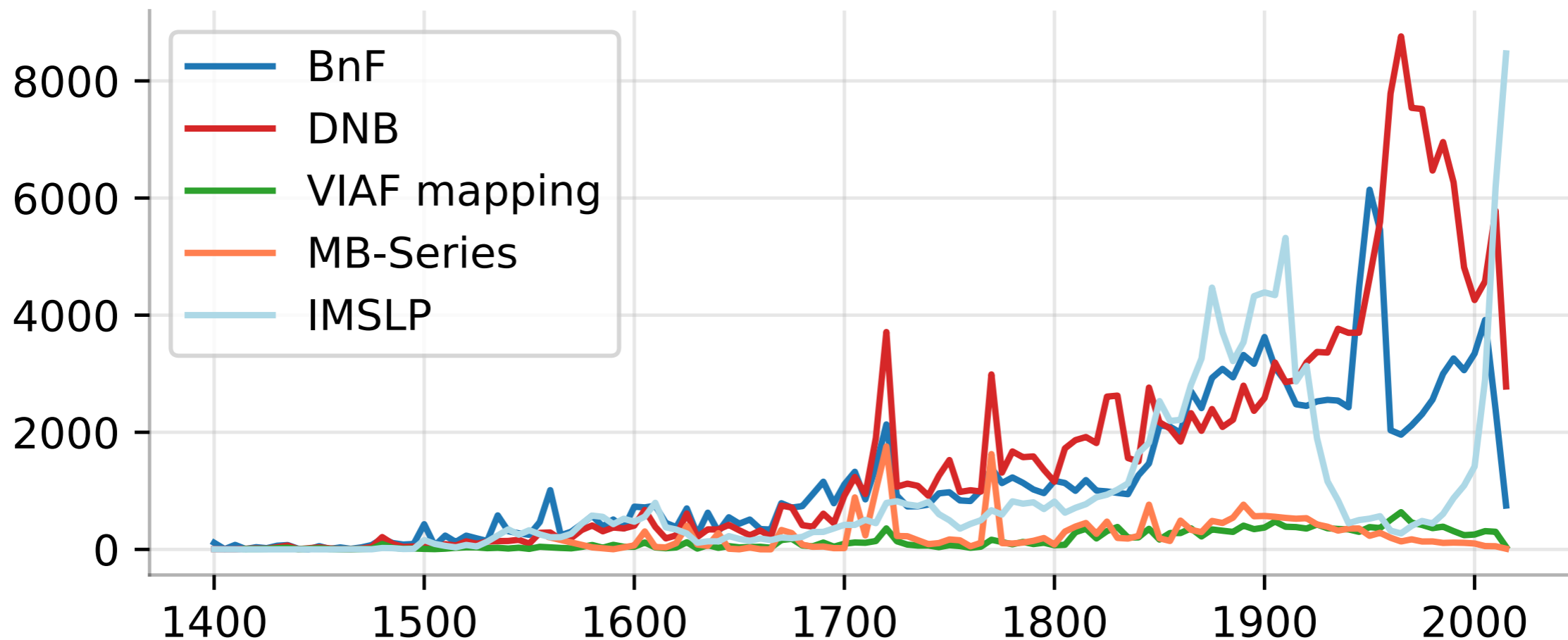| Library	BNF only	BNF and DNB	DNB only
Library of Congress	11585	14987	26493
Bibl. Nac. de España	755	2318	1471
Nat. Lib. of Australia	496	2527	1287
Nat. Lib. of Israel	505	2306	1171
Nat. Lib. of Poland	228	1516	1197
Lib. and Arch. Canada	321	898	578
Bibl. nat. du Québec	203	501	401
Nat. Lib. of Catalonia	96	642	323

0% 20% 40% 60% 80% 100%

Number of works from the BnF and DNB mapped through VIAF to the 8 libraries
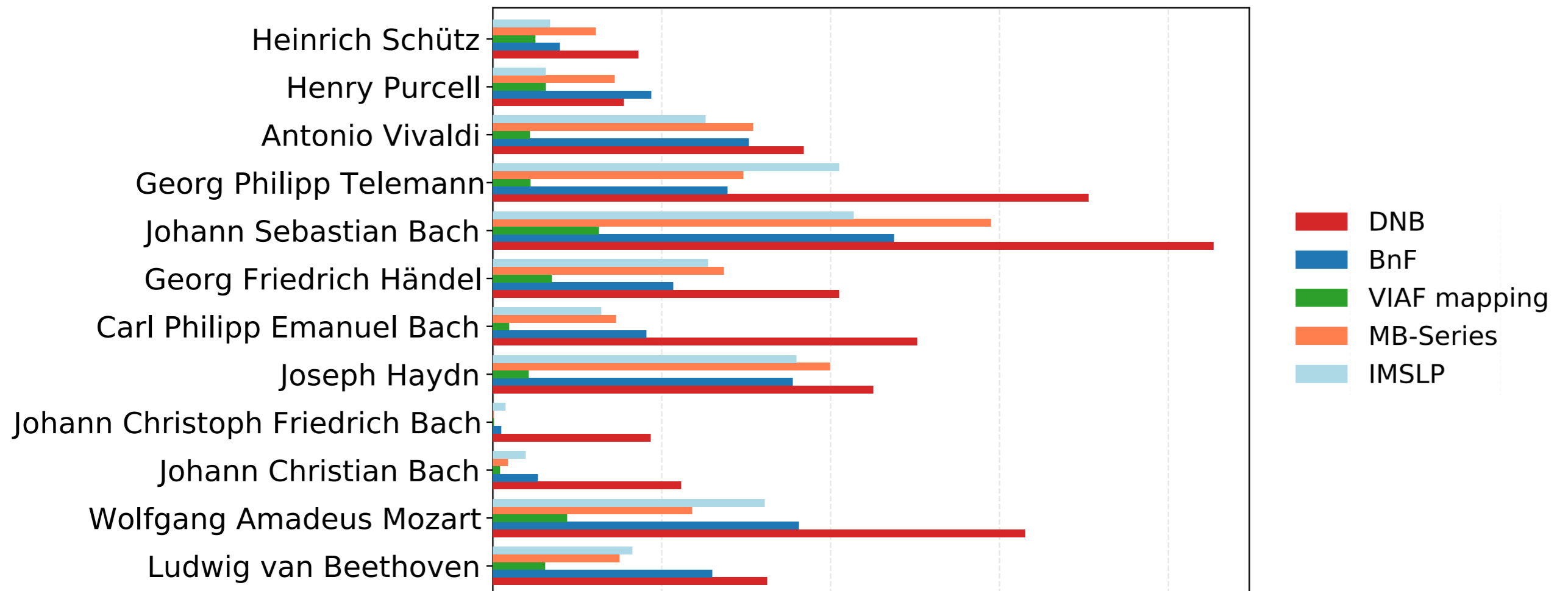with the highest number of links

Variation over time

- ✤ DNB gnd:dateOfProduction (35%)
- ✤ BnF dcterms:date or bnf-onto:firstYear (63%)
- ✤ DNB and BnF composer birth / death years to approximate works without date
- ✤ IMSLP composition date, first publication, first performance

Works (1400 - 2020)

Composers with most work

✤ Mapping through VIAF for these is higher

  ✦ 30% from BnF / 18% from DNB

✤ Systematically more works by composer in DNB than in other datasets

  ✦ 1.65 times more than in BnF

  ✦ 2.0 times more than in IMSLP and MB-Series

Bach example

✤ DNB gnd:broaderTermPartitive / BnF ore:isAggregatedBy

✤ Up to three levels in both DNB and BnF

  ✦ Example: *Das wohltemperierte Klavier* in DNB

    ✤ Preludes and fugues (though not all, only 37 for book I) listed as distinct works

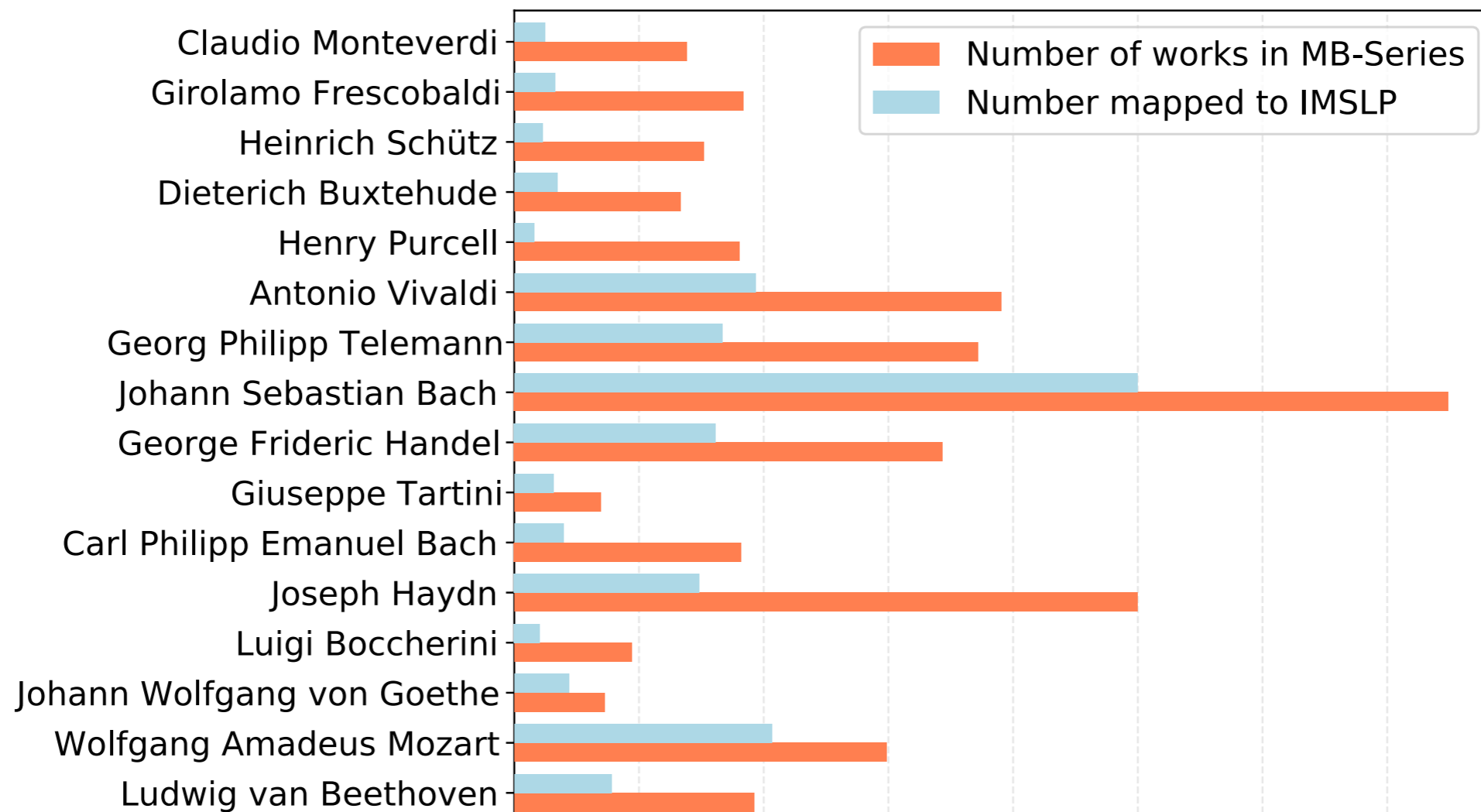✤ Top-level work entities are significantly better mapped in V than lower ones

	DNB	BnF mapping		LC mapping	
		#	%	#	%
Total	2,133	311	15%	625	29%
Top-level	744	232	**31%**	268	**36%**
2nd level	1,389	79	6%	357	26%
3rd level	280	3	1%	22	8%

✤ Mapping also differers significantly when looking at the me of performance

Claudio
Girolamo
Hein
Dieterich
H
Ant
Georg Philip
Johann Seb
George Fric
Gius
Carl Philipp Em
Jo
Luig
Johann Wolfgang
Wolfgang Amac
Ludwig van
Carl Maria
Fra
He
Felix M
Rober

Joh
Johan

RISM | DIGITAL CENTER

Mapping with MB, IMSLP and Wikidata

✤ Wikidata lists about 10,000 music compositions (and 1,000 musical works)

✤ Of the 24'198 works listed in MB-Series, 9,079 have a link to IMSLP

	IMSLP		MB-Series		Wikidata	
	#	%	#	%	#	%
DNB	1,303	0.8%	106	0.4%	509	4.8%
BnF	2,006	**1.2%**	173	0.7%	689	6.5%
LC	1,338	0.8%	31	0.1%	568	5.4%
VIAF	1,745	1.1%	1,100	**4.5%**	702	**6.7%**

Summary

✤ Number of works by composers in all library datasets are very low

    ✦ Large datasets contain only a couple of works for most composers

✤ Mapping with VIAF remains relatively low

    ✦ Mapping between DNB and LC is the highest

✤ Looking over time

    ✦ BnF contains more works than DNB until the 17th century

    ✦ Curves for each dataset can fluctuate over short periods

✤ The definition of a work varies significantly between datasets

    ✦ DNB has systematically a lower-level granularity with movements listed as distinct works

RISM | DIGITAL CENTER

Thank you!

laurent.pugin@rism.digital

https://rism.digital

RISM | DIGITAL CENTER