Geographic differences in ICT research

An analysis of keywords and key-phrases

Alex Surpatean, Ad Notten, Bulat Sanditov and Jojo Jacob

Research Questions

R&D performance (productivity & impact) can be improved by acquiring more and better knowledge. The problem is, however, that knowledge, despite having all the characteristics of a public good, does not behave as such. Its diffusion is slow and geographically localized.

Therefore:
- Does the topical focus of ICT research in the EU differ from the R&D topics in other regions and countries?

The European Research Area's aim is breaking down the "barriers" to knowledge diffusion in order to create a genuine single market for knowledge, research and innovation.

As such:
- Has the ERA succeeded in bringing the ICT research efforts of the EU member states closer together or are there still differences?

UNITED NATIONS
UNIVERSITY

UNU-MERIT

Bibliometrics

- Co-publications, and as such co-authorships, are well established sources for metrics used in the construction of research performance indicators. Publication and patent records offer a wealth of data beyond the obvious.

- Next to citation analysis, used in research performance (productivity and impact) measurement, co-word analysis offers an important alternative methodology for Researching social, contextual and methodological differences concerning the R&D performed .

- It is important to understand how collaborative ties are formed, not only at the country-level, but also at the sectoral or disciplinary level. In order to understand the linkage patterns of bi/multi-lateral cooperation, it is necessary to observe collaborative linkages *between* the different R&D performers in order to analyse how sectors and disciplines in different countries or groups work together (domain & social knowledge).

Data used

- Yearly surveys by DG CONNECT targeting project coordinators.

- Elsevier Scopus custom data set focusing on the same source titles as those published in by EU funded researchers.

- Stratified samples based on year and publication type.

Population	Control Population	Confidence Level	Confidence Interval	Control Sample
30996132	6846926	0.99	0.02	4158
Population	Focal Population	Confidence Level	Confidence Interval	Focal Sample
24664	19273	0.99	0.02	3422

Distribution of publication output over the affiliation countries, focal vs. control

	Focal			Control	
Percentage	Publications	Country	Country	Publication	Percentage
12.3	763	Germany	United States	1064	19.0
11.7	723	Italy	China	591	10.5
9.7	604	United Kingdom	Germany	412	7.4
8.8	548	Spain	United Kingdom	358	6.4
6.8	423	France	France	307	5.5
4.8	298	United States	Japan	293	5.2
4.4	275	Switzerland	Italy	214	3.8
4.3	268	Netherlands	Canada	193	3.4
4.2	260	Greece	Spain	156	2.8
3.1	195	Sweden	Australia	144	2.6
3.1	191	Austria	India	132	2.4
2.9	180	Belgium	South Korea	107	1.9
1.7	104	China	Switzerland	95	1.7
1.7	104	Portugal	Netherlands	92	1.6
1.3	80	Finland	Taiwan	86	1.5
1.2	77	Ireland	Russian Federation	82	1.5
1.2	77	Israel	Sweden	76	1.4
1.2	73	Poland	Poland	67	1.2
1.1	69	Denmark	Austria	66	1.2
0.9	55	Norway	Brazil	66	1.2
0.8	49	Canada	Singapore	61	1.1

Econometric models – negbin models results

	1	2	3	4	5	6	7	8
Project duration	0.046***	0.045***	0.024***	0.025***	0.058***	0.050***	0.031***	0.026***
	(0.012)	(0.012)	(0.008)	(0.008)	(0.012)	(0.012)	(0.008)	(0.008)
log(Total funding)	0.807***	0.509***	0.446***	0.131				
	(0.091)	(0.116)	(0.142)	(0.154)				
Number of participating organizations					0.063***	0.030**	0.021***	0.010*
					(0.009)	(0.013)	(0.006)	(0.006)
log(Funding per participants)					0.506***	0.408***	-0.149	-0.027
					(0.120)	(0.123)	(0.143)	(0.156)
Patents	1.114***	1.130***	1.102***	1.113***	1.161***	1.141***	1.132***	1.123***
	(0.112)	(0.110)	(0.112)	(0.109)	(0.111)	(0.109)	(0.112)	(0.110)
University share	2.212***	2.199***	2.039***	2.026***	1.925***	2.064***	1.887***	1.982***
	(0.242)	(0.229)	(0.207)	(0.203)	(0.219)	(0.219)	(0.195)	(0.199)
Diversity of national science base		3.603***		3.295***		4.237***		3.067***
		(1.007)		(0.871)		(1.114)		(0.889)
CP-FP-INFSO			1.103***	1.259***			1.723***	1.439***
			(0.282)	(0.281)			(0.280)	(0.302)
CP-IP			1.280***	1.489***			2.212***	1.739***
			(0.383)	(0.387)			(0.334)	(0.383)
NoE			3.150***	3.120***			3.603***	3.231***
			(0.550)	(0.555)			(0.561)	(0.560)
CP-CSA-Infra_PP			1.127***	1.365***			1.468***	1.354***
			(0.421)	(0.471)			(0.397)	(0.482)
Year dummies	included	included	included	included	included	included	included	included
Constant	-14.219***	-12.813***	-9.235***	-7.474***	-9.648***	-11.373***	-1.717	-5.308**
	(1.217)	(1.240)	(1.860)	(1.960)	(1.379)	(1.655)	(1.598)	(2.078)
Observations	14,664	14,664	14,664	14,664	14,664	14,664	14,664	14,664

Ontologies used indexing the FP7 funded ICT research keyword data

- EI thesaurus (engineering, technology, physical sciences)
- Emtree medical terms (life sciences & health sciences)
- MeSH (life sciences & health science)
- Geobase Subject Index (geology, geography, earth and environmental science)
- FLX terms, WTA terms (fluid sciences & textile sciences)
- Regional Index (geology, geography, earth and environmental science)
- Species Index (biology, life sciences)

Probabilistic Affinity Index

Probabilistic Affinity Index (PAI): the ratio of the observed number of links (in the focal sample) to expected one (in the control sample), indicates relative tendency of co-authorship:

- PAI = n_{ij} / E(n_{ij}) = $n_{..}$ * n_{ij} / ($n_{i.}$ * $n_{.j}$).

Where E(n_{ij}) is expected value of n_{ij} and $n_{..}$ Is the sum of all cells of the co-occurrence matrix (diagonal discarded).
Here, a value of more than 1 indicates that there are more cooperative links than were expected, and vice versa, below 1 indicates less links than expected. We then normalise the index to get a value between -1 to 1:

- Normalized PAI = $(1 - PAI^2) / (1 + PAI^2)$.

Normalized probabilistic affinity index, regional comparison

	EU28	c_EU28	c_US	c_JP	c_KR	c_BRICS
EU28		0.366	0.220	0.082	0.004	0.205
c_EU28	0.366		0.326	0.074	-0.027	0.227
c_US	0.220	0.326		0.120	0.122	0.241
c_JP	0.082	0.074	0.120		0.336	0.028
c_KR	0.004	-0.027	0.122	0.336		-0.028
c_BRICS	0.205	0.227	0.241	0.028	-0.028	

Comparing the focal EU-28 sample with various regional control samples

UNITED NATIONS UNIVERSITY

UNU-MERIT

Normalized probabilistic affinity index, national comparison (intra-EU)

	AT	BE	BG	CY	CZ	DE	DK	EE	ES	FI	FR	UK	GR	HR
AT		0.08	0.08	-0.02	0.08	0.18	0.07	0.01	0.00	0.00	0.03	0.04	0.01	-0.12
BE	0.08		0.01	0.05	-0.09	0.08	0.04	-0.04	0.07	0.00	0.14	0.02	-0.01	-0.16
BG	0.08	0.01		0.56	0.31	-0.15	-0.04	0.79	-0.15	0.31	-0.34	-0.24	0.14	0.76
CY	-0.02	0.05	0.56		0.15	-0.25	0.10	0.67	-0.13	0.10	-0.23	0.02	0.21	0.46
CZ	0.08	-0.09	0.31	0.15		-0.08	0.14	0.57	-0.04	0.11	-0.03	-0.08	-0.10	0.59
DE	0.18	0.08	-0.15	-0.25	-0.08		-0.02	-0.39	0.19	0.01	0.20	0.24	0.10	-0.34
DK	0.07	0.04	-0.04	0.10	0.14	-0.02		-0.10	-0.02	0.07	0.03	0.01	-0.02	0.14
EE	0.01	-0.04	0.79	0.67	0.57	-0.39	-0.10		-0.28	-0.26	-0.28	-0.34	-0.34	0.89
ES	0.00	0.07	-0.15	-0.13	-0.04	0.19	-0.02	-0.28		-0.06	0.18	0.18	0.09	-0.20
FI	0.00	0.00	0.31	0.10	0.11	0.01	0.07	-0.26	-0.06		0.00	0.02	0.06	0.09
FR	0.03	0.14	-0.34	-0.23	-0.03	0.20	0.03	-0.28	0.18	0.00		0.15	0.06	-0.23
UK	0.04	0.02	-0.24	0.02	-0.08	0.24	0.01	-0.34	0.18	0.02	0.15		0.13	-0.31
GR	0.01	-0.01	0.14	0.21	-0.10	0.10	-0.02	-0.34	0.09	0.06	0.06	0.13		0.03
HR	-0.12	-0.16	0.76	0.46	0.59	-0.34	0.14	0.89	-0.20	0.09	-0.23	-0.31	0.03	
HU	0.00	-0.04	0.32	0.27	0.23	-0.14	0.06	0.70	-0.03	0.04	-0.12	-0.10	-0.03	0.63
IE	0.02	0.04	0.23	0.24	0.01	0.04	-0.05	0.28	-0.02	0.24	-0.08	0.05	0.06	0.09
IT	0.05	0.05	-0.30	-0.20	-0.16	0.24	0.03	-0.34	0.20	-0.05	0.22	0.22	0.09	-0.23
LT	0.03	0.06	0.69	0.54	0.20	-0.23	0.13	0.78	-0.12	0.13	-0.11	-0.15	-0.06	0.47
LU	0.04	0.08	0.38	0.73	0.03	-0.36	-0.10	0.90	-0.14	0.11	-0.26	-0.15	0.00	0.62
LV	0.17	0.14	0.49	0.28	0.44	-0.44	0.21	0.76	-0.32	0.01	-0.03	-0.42	-0.14	0.61
MT	-0.48	-0.59	-0.26	0.12	0.89	-0.79	-0.64	1.00	-0.72	-0.41	-0.53	-0.46	-0.63	0.98
NL	0.04	0.16	-0.19	-0.12	0.00	0.16	0.05	-0.02	0.06	-0.02	0.09	0.11	0.01	-0.14
PL	0.02	0.14	0.47	0.36	0.30	-0.07	0.13	0.30	0.05	0.03	-0.02	-0.08	0.03	0.39
PT	-0.03	-0.06	0.16	0.34	0.01	-0.06	0.00	0.25	0.04	0.07	-0.02	0.01	0.14	0.07
RO	0.01	-0.10	0.52	0.31	0.31	-0.17	0.03	0.68	-0.13	0.15	-0.09	-0.10	0.05	0.62
SE	0.07	0.07	-0.17	-0.16	0.06	0.11	0.10	0.07	0.06	0.23	0.06	0.04	-0.01	0.19
SI	0.11	-0.09	0.31	0.24	0.21	-0.18	0.14	0.11	-0.08	0.20	-0.21	-0.09	0.06	0.49
SK	0.30	-0.18	0.72	0.32	0.73	-0.20	0.17	0.11	-0.18	0.13	-0.28	-0.24	-0.11	0.63
	AT	BE	BG	CY	CZ	DE	DK	EE	ES	FI	FR	UK	GR	HR

Normalized probabilistic affinity index, national comparison (intra-EU)

	HU	IE	IT	LT	LU	LV	MT	NL	PL	PT	RO	SE	SI	SK
AT	0.00	0.02	0.05	0.03	0.04	0.17	-0.48	0.04	0.02	-0.03	0.01	0.07	0.11	0.30
BE	-0.04	0.04	0.05	0.06	0.08	0.14	-0.59	0.16	0.14	-0.06	-0.10	0.07	-0.09	-0.18
BG	0.32	0.23	-0.30	0.69	0.38	0.49	-0.26	-0.19	0.47	0.16	0.52	-0.17	0.31	0.72
CY	0.27	0.24	-0.20	0.54	0.73	0.28	0.12	-0.12	0.36	0.34	0.31	-0.16	0.24	0.32
CZ	0.23	0.01	-0.16	0.20	0.03	0.44	0.89	0.00	0.30	0.01	0.31	0.06	0.21	0.73
DE	-0.14	0.04	0.24	-0.23	-0.36	-0.44	-0.79	0.16	-0.07	-0.06	-0.17	0.11	-0.18	-0.20
DK	0.06	-0.05	0.03	0.13	-0.10	0.21	-0.64	0.05	0.13	0.00	0.03	0.10	0.14	0.17
EE	0.70	0.28	-0.34	0.78	0.90	0.76	1.00	-0.02	0.30	0.25	0.68	0.07	0.11	0.11
ES	-0.03	-0.02	0.20	-0.12	-0.14	-0.32	-0.72	0.06	0.05	0.04	-0.13	0.06	-0.08	-0.18
FI	0.04	0.24	-0.05	0.13	0.11	0.01	-0.41	-0.02	0.03	0.07	0.15	0.23	0.20	0.13
FR	-0.12	-0.08	0.22	-0.11	-0.26	-0.03	-0.53	0.09	-0.02	-0.02	-0.09	0.06	-0.21	-0.28
UK	-0.10	0.05	0.22	-0.15	-0.15	-0.42	-0.46	0.11	-0.08	0.01	-0.10	0.04	-0.09	-0.24
GR	-0.03	0.06	0.09	-0.06	0.00	-0.14	-0.63	0.01	0.03	0.14	0.05	-0.01	0.06	-0.11
HR	0.63	0.09	-0.23	0.47	0.62	0.61	0.98	-0.14	0.39	0.07	0.62	0.19	0.49	0.63
HU		0.08	-0.09	0.38	0.29	0.22	0.91	-0.04	0.19	0.21	0.45	0.07	0.37	0.38
IE	0.08		-0.09	0.21	0.45	0.26	-0.32	-0.06	0.22	0.07	0.17	-0.02	0.06	0.41
IT	-0.09	-0.09		-0.24	-0.14	-0.02	-0.48	0.13	-0.12	0.04	-0.13	0.04	-0.08	-0.19
LT	0.38	0.21	-0.24		0.30	0.77	0.43	-0.13	0.50	0.07	0.32	-0.21	0.65	0.56
LU	0.29	0.45	-0.14	0.30		0.76	0.05	-0.08	0.35	0.05	0.55	-0.01	0.26	0.40
LV	0.22	0.26	-0.02	0.77	0.76		0.42	-0.30	0.34	0.27	0.49	-0.03	0.58	0.73
MT	0.91	-0.32	-0.48	0.43	0.05	0.42		-0.26	0.07	-0.29	0.92	0.50	0.51	0.12
NL	-0.04	-0.06	0.13	-0.13	-0.08	-0.30	-0.26		-0.16	0.08	0.00	0.07	0.03	-0.19
PL	0.19	0.22	-0.12	0.50	0.35	0.34	0.07	-0.16		-0.07	0.21	-0.01	0.03	0.43
PT	0.21	0.07	0.04	0.07	0.05	0.27	-0.29	0.08	-0.07		0.21	0.03	0.35	-0.05
RO	0.45	0.17	-0.13	0.32	0.55	0.49	0.92	0.00	0.21	0.21		-0.01	0.16	0.47
SE	0.07	-0.02	0.04	-0.21	-0.01	-0.03	0.50	0.07	-0.01	0.03	-0.01		-0.11	-0.16
SI	0.37	0.06	-0.08	0.65	0.26	0.58	0.51	0.03	0.03	0.35	0.16	-0.11		0.50
SK	0.38	0.41	-0.19	0.56	0.40	0.73	0.12	-0.19	0.43	-0.05	0.47	-0.16	0.50	

Conclusion and future directions

- We measured the degree of overlapping national and regional research interests by analysing the keywords that describe the publication output. This was presented as a measure of overlapping research interests, but not computed based on explicit cooperation ties, but through indirect labels of interest.

- We noted that the DG CONNECT funded output has a positive affinity with all the considered control sample groups, with affinity greater than the baseline expected value given random independent occurrence.

- The DG CONNECT sponsored output has greatest affinity with the EU-28 control group of non-DG CONNECT-funded projects, which is to be expected. The focal sample also has a large positive affinity with the US and BRICS control sample groups, although slightly lower than the EU-28 control has with these groups

Conclusion and future directions

• The focal sample has lesser affinity with the Japanese and South Korean control samples, and we notice that the affinity of the EU-28 focal with these samples is greater than that of the EU-28 control with these samples.

• The EU has been shown to own a large and high quality knowledge pool in the field of ICT which, stimulated by a high degree of internationalization, which should grow further in the years to come. Attracting and educating more knowledge workers will solidify the EU leadership position in ICT research. A further step could be engaging in further, and perhaps less regulated, collaborations with new research partners in- and out-side of the EU. As shown, there seems to be room for further knowledge exchange with Eastern European partners as well as with BRICS countries and third countries such as Japan and South Korea. It is inevitable that knowledge exchanges with these countries will lead to further innovations and a strengthening of the EU's "Excellence in Science" – an H2020 goal.

SmartEIZ Conference

Smart Specialization Strategies
Contemporary Challenges of its Design and Implementation

25-26 September, 2018
The Institute of Economics, Zagreb

Thank you!