

Improving the collection of data on Social Science and Humanities. Experiences from the South

Michael Kahn, Centre for Research and Evaluation of Science and Technology, Stellenbosch University, Matieland, Stellenbosch, South Africa.
mikejkahn@gmail.com

Regina Gusmão, Ministry of Science and Technology, Brasilia, Brazil.
regina.gusmao@mct.gov.br
CODATA 2010
26 October 2010

1. New profile of SSH – MDGs; global dynamics; ICT revolution/rise of the services economy. Move to 'knowledge societies?'
2. The problem of collecting data on SSH research problems
 - Incompatibilities and lack of articulation of databases
 - Discipline boundaries, sector definitions, measurement of headcounts and full-time equivalents and the outputs of SSH research
 - Understanding the employability, occupations and sectoral distribution of SSH masters and doctoral degree holders.
3. SSH information systems: cases - South Africa and Brazil
4. Minimum data set & standardization

- wo important milestones in 2010
- *UNESCO/ISSC ≈ World Social Science Report 2010: Knowledge Divides*
- *OECD ≈ The OECD Innovation Strategy – Getting a head start on tomorrow*
- rocesses and publications in turbulent times
- ace one another across an old divide

“... Social science expertise is in high demand by policy makers, media and the public.... But with success and growth have come criticism. It has been pointed out that few economists foresaw the economic crisis that started in 2008 and that conflicting advice has been given on dealing with it.”

(WSSR, 2010. *Summary. World Social Science Report 2010. Knowledge Divides*. Paris: UNESCO ISSC)

“ The 2008-09 economic downturn has led to reduced potential output growth, rising unemployment and soaring public debt. To recover, countries need to find new and sustainable sources of growth Many countries have stagnating or declining populations and face diminishing returns from labour inputs and investment in physical capital. Future growth must therefore come from innovation-induced productivity growth.

(OECD2010. *Fostering Innovation to Strengthen Growth and Address Global and Social Challenges. Executive Summary*. Paris: OECD)

“... what went wrong was innovation in financial services which resulted in the release of attractive new products to the market. They diffused rapidly and widely and then lost value. The rest is history, a painful history for those who lost homes, savings and businesses. As the first signs of recovery appear, the question being asked is whether this can happen again?”

(Gault, F 2010. *Innovation Strategies for a Global Economy. Development, implementation, measurement and management* Cheltenham: Edward Elgar)

- ... “(One notes) the scarcity of data needed for the comparison of research capacities and for the assessment of strategies in different parts of the world, especially in the social sciences. There is an urgent need for data-gathering to support these comparisons and analyses.” (WSSR, 2010)
- But measurement is theory laden - R&D linked to economic growth; what is measured is deemed to have economic value; how do we ensure commensurability?

Toward a science of science policy

- ‘Relating R&D to innovation in any but a general way is a tall order, but not a hopeless one. We need econometric models that encompass enough variables in a sufficient number of countries to produce reasonable simulations of the effect of specific policy choices. This need won’t be satisfied by a few grants or workshops, but demands the attention of a specialist scholarly community. As more economists and social scientists turn to these issues, the effectiveness of science policy will grow, and of science advocacy too’ (Marburger, 2005: 1087).
- Couzens: contested! It already exists

World Social Science Report 1999

- Covered 29 OECD member states
- Sought to quantify scale of inputs; graduates; Frascati Manual R&D definition and collection guidelines
- SS R&D financials for 21 countries; personnel data for 19 – excludes large systems: FR, US, UK, KR, and smaller - CZ, GR, IT, LX, NZ, SL.
- Commentary on the practice of SS; notes ‘Data ... often not very accessible to researchers’; identifies blurring of SS/H boundary as problem

World Social Science Report 2010

- OECD 30; EU-27 + other large performers of R&D
- Expenditure on R&D; development indicators
- Demographics of researchers
- Student enrolment and graduation
- Scientific output - publications
- Main sources: OECD, Eurostat, UNESCO Institute for Statistics. RICYT (Latin America)
- Other: national statistical organizations; commercial databases

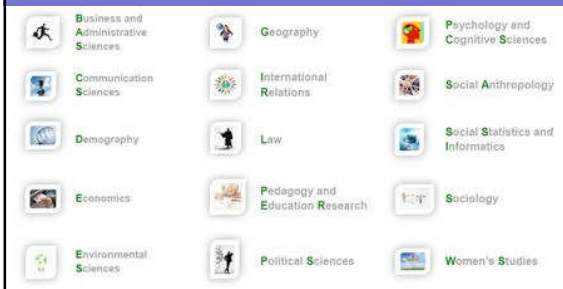
What counts as R&D

- Frascati 6th edition : ‘creative work undertaken on a systematic basis in order to increase the stock of knowledge, including **knowledge of man, culture and society**, and the use of this stock of knowledge to devise new applications’
- Recognizes increasing role of Services sector
- Basic, applied research, experimental development
- Related scientific and technical activities (STA) excluded
- Clinical trials; software development ... IN
- Much Social Science activity ... OUT

Problems

- Business R&D difficult to trace → purposive survey
- SS often excluded from BERD: “designed out”
 - SS important in Services; > 60% GDP
 - R&D in Services 40% BERD in AU; 25% UK
 - R&D tax incentives may exclude SS
- University and Gov R&D by census;
- NPO purposive – boundaries porous
- FTE “must” be < 100%; estimation; diary studies
- Doctoral, postdocs, research assistants, masters – FTE?

Classification: ESF OECD agree;
UNESCO education not SS; national 'do it our way'



	1999	2000	2001	2002	2003	2004	2005	2006
Australia		38507		42780		47417		50934
Austria				8577		8281		8944
Belgium	11873	11778	12034	12366	12388	13548	13383	14417
Canada	33020	33300	34200	34910	38900	41380	43420	-
Czech Republic	3301	3789	4249	4383	4318	4771	7576	8352
Denmark	5722	5556	6106	7373	7689	7846	8242	8763
Finland	10555	10999	11008	12391	13033	13037	12879	12849
France	56717	61583	62427	63555	64403	65488	66200	67935
Germany	58698	67087	67382	71302	68343	69764	85383	86933
Greece	10471		8544		9072		11356	12110
Hungary	4768	5852	5938	5999	5957	5902	5911	6073
Iceland	489		515		502		585	563
Ireland	2285	2148	2479	2797	3474	4151	4400	4672
Italy	25209	25696	27146	28301	27774	28226	37073	37638
Japan	178418	179116	200272	170512	172396	177421	180498	184319
Korea	21123	23874	23083	24383	26419	25922	27416	28368
Luxembourg		22	30		30	143	157	159
Mexico	10648				17135	16043	16691	
Netherlands	12491	15480	15750	10448	10211	10545	10661	10931
New Zealand	4996		8655		9955		11731	
Norway	5521		5870		6281	6800	7512	7670
Poland	38264	34248	36597	37275	38469	39716	40489	37653
Portugal	8242	8592	8942	9502	10062	10509	10956	12026
Slovak Republic	4254	5000	4891	4609	5273	6509	6488	7370
Spain	33840	42064	46964	45727	49196	51616	54028	55443
Sweden	14623		15851		17146	17794	15125	14740
Switzerland		8428		11340		12338		12710
Turkey	14621	16992	16798	17544	24226	24742	25434	26713
United Kingdom							67719	69499
United States	186049							

ARTICLES	Web of Science			SCOPUS	
	SCH	SSCI	A&HCI	SOCSCI	ARTS
Bulgaria	1586	33	5	83	6
Estonia	696	86	8	91	14
Hungary	3658	172	43	309	70
Lithua	229	16	0	12	0
Lithuania	810	64	54	177	37
Poland	10615	258	75	426	44
Romania	2062	69	50	97	29
Slovak Republic	1049	108	71	159	59
Slovenia	1833	137	39	343	20
Austria	7267	525	84	614	57
Belgium	10484	1158	254	1263	130
Canada	38763	5861	1094	5719	479
Cyprus	289	68	13	114	4
Czech Republic	5116	263	86	302	25
Denmark	7975	833	78	783	59
Finland	7078	891	87	963	69
France	42563	2200	1018	2872	398
Germany	59628	4678	924	4651	438
Greece	7320	457	84	738	65
Ireland	5045	754	146	592	48
Italy	33365	1758	362	2214	181
Luxembourg	176	21	1	33	1
Malta	60	10	4	9	1
Netherlands	18772	3573	316	3559	194
Portugal	4938	289	33	463	26
Spain	27338	2298	518	2519	193
Sweden	14381	1860	131	1816	116
United Kingdom	51844	12786	13732	13732	1450
TOTAL	381068	21189	7084	22295	2211

The limits of EU law

- Decision № 1608/2003/EC of the European Parliament and of the Council for statistics on S&T.
- Legal implementation measure 753/2004

FOS	Sectors	GERD	R&D PERSONNEL	RESEARCHERS
1-DIGIT FOS	GOV, HES	Optional (R&D Regulation) Biennially	HC: Obligatory (R&D Regulation) by sex; four yearly FTE: Optional (R&D Regulation) by sex; biennially	
	TOTAL, BES, PNP	Voluntary	HC and FTE: Voluntary; by sex	
2-DIGITS FOS	TOTAL, GOV, HES, PNP	Voluntary	-	HC: Voluntary; by sex

What we know & what we know we don't know

- 25: Czech Republic, Hungary, Poland, Romania, Russian Federation, the Slovak Republic, Slovenia, Turkey, Mexico, Chinese Taipei, Japan, Singapore, Austria, Belgium, Canada, Denmark, Germany, Ireland, Italy, Luxembourg, Netherlands, Norway, Portugal, Spain, and Sweden
- 17 EU; 20 OECD
- Of top five R&D performing nations, only DE and JP have official data

The need for estimates

- Large performers missing: FR, UK, also FI.
- Newcomers more compliant? Why?

		BUS		HE		Gov		
		SET	Tot	SET	SSH	Tot	SET	SSH
France	HC	107401	100849	70998	29851	31936	27146	4790
(2003)	FTE	100646	59047	43695	15352	31936	27146	4790
UK*	HC	241127	139099	102028	9893	8962	932	
(2006)	FTE	95592	67719	39059	28660	9311	8563	748
USA	HC			297000	275 000			
(2006)	FTE			120 000	111 000			
China	HC			-	-			
(2005)	FTE			166 400	55508			
Russia	HC	221445	30111	26130	3981	139378	126413	13235
(2005)	FTE	237959	70494	61595	8899	154827	140425	14402

South Africa

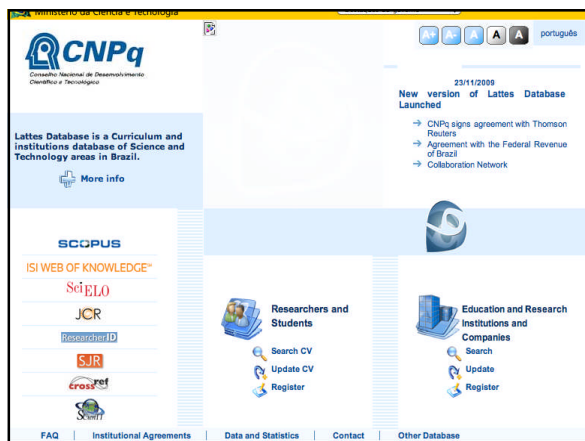
- Small system of innovation with new developmental goals
- Many data sets; design-specific; gaps; uncoordinated
 - Higher Education Management Information System (HEMIS):
 - Audited
 - CESH system of classification
 - Research output database
 - National R&D Survey (OECD – compliant)
 - SA Knowledgebase (private)
 - NRF individual, not group rating
 - State lab KPIs under-utilized
 - National Student Financial Aid Scheme data underutilized
 - M & D database absent
 - Research Information Management System (WIP)

- Future HEPMS under discussion
 - Resistance to measurement; overloaded; information demands; data for compliance rather than change management purposes
 - Ranking: absolute or relative?
 - How to account for community engagement?
 - SATN performance indicators db being populated
 - Skills to operate MIS?
 - Are students 'customers'?
- Data generated on public funds should be in the public domain c.f. *OECD Principles*

Brazil

- Brazil as a world power –military regime early 1970s;
- “Modernizing” agenda: federal, state and private HEIs.
- Central planning : National Postgraduate Programmes (PNPG) from 1974
- Agencies
 - CNPq (S&T research)
 - CAPES (HR for research)
- University and department rating system
- Own classification system:
 - Social Sciences: administration, architecture and urban studies, urban planning, information sciences, communications, law, demography, economics, social services and tourism
 - Humanities: anthropology, archaeology, political science, education, philosophy, geography, history, psychology, sociology and theology. Languages, literature, arts excluded

- M and PhD programmes (2008) now 2,568
- M and PhD graduation up threefold
- SSH 1/4 of the 10, 000 PhDs awarded in 2008
- Research groups increase 5 x 1993 to 2008.
- 2000 and 2008, SS research groups 2,600 to nearly 7,000
- 66% Ph.Ds graduates in 2008 employed HEIs; 18% in public administration, national defence or social security; only 1.2% in manufacturing industries.
- Excellent information – how?



- Quantification through *Plataforma Lattes* of CNPq – unique cv database that allows tracking of researchers & research groups. Full demographic query system
- Builds on national system engineering skills developed from 1980s onward during import substitution era
- Belgium: ECOOM
- South Africa: SA Knowledgebase some

We are not alone

- Numerous examples of MIS that capture data on HEIs, student population, university achievement
- Bibliographic databases (Scholarly book series; electronic distribution) – much WIP e.g. ERIH; ESF-SCH bibliometric database
- Institutional profiles e.g. UniBasel emerging
- Country profiles e.g. Portal of Science Journals of Croatia
- It's competitive: THES and QS have parted company ...
- Many projects on University ranking EUMEDA; U-RANK
- Input side - national systems; output side - private



Centre for Research on
Evaluation, Science and Technology

Comments on the data

- Balance between centralization/decentralization' and standing of science in government and in national development
- What is underlying purpose of the data collection?
- Access to data is pivotal: what is possible in Brazil or France or South Africa may be impossible elsewhere
- FTE of Researchers is patchy; costly; estimates contested
- Student data more robust
- What counts as a publication?
- Propensity to publish varies by discipline
- Mobility data is weak



Centre for Research on
Evaluation, Science and Technology

Toward the Future

- What must we know on research capacity in SSH?
- Who wishes to know?
- What is available and who holds the data? Privacy issues.
- What must be collected rather than what is 'nice to have'?
- What constitutes a minimum data set for international benchmarking? Feasibility?
- What effort needed to obtain more complete input data? What are the restrictions on data access?
- Prospects for more complete bibliometric data?
- Roles for ISSC, ESF, OECD, UNESCO, Eurostat especially in relation to WSSR?
- Role for CODATA?



Centre for Research on
Evaluation, Science and Technology