Lessons from and for research policy impact evaluation

Stefan Kuhlmann, University of Twente
Science, Technology, & Policy Studies
(STəPS)

EUFORDIA Conference, Prague, 24-25 February 2009

Overview

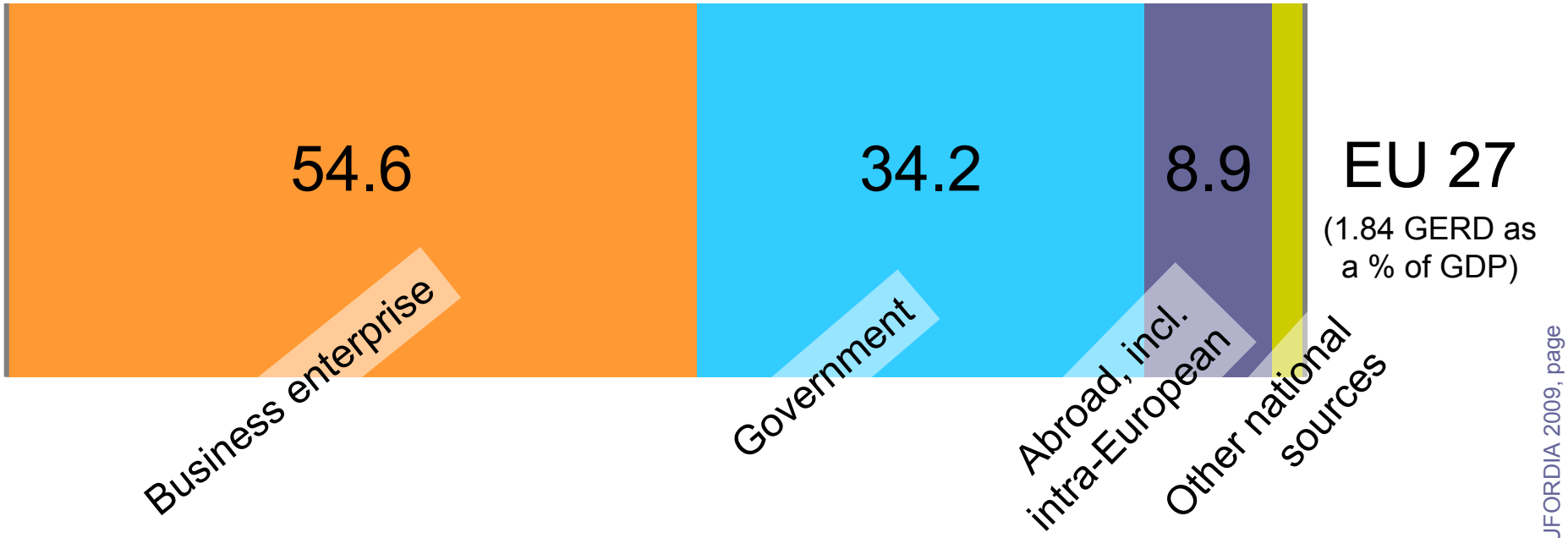
1. Purposes, options, limitations of R&D policy impact evaluation
2. R&D policy and impact evaluation for Europe
   - Strengths and weaknesses of FP evaluation practice
   - New perspectives: Knowledge dynamics, stakeholders, ERA
3. Outlook: Strategic evaluation

Overview

1. **Purposes, options, limitations of R&D policy impact evaluation**
2. R&D policy and impact evaluation for Europe
3. Outlook: Strategic evaluation

Considerable public R&D investment

R&D expenditure in Europe by main sources of funds (%), 2006



Source: Science, Technology and Competitiveness key figures report 2008/2009

Evaluation of public investments:

Efficiency & usefulness (legitimisation); lessons learned

Varieties of R&D (policy) evaluation

	Grants for open-ended research	***Grants and programs for strategic research (e.g. EU FPs)***	***Contracts, mission-oriented R&D programs***
Ex-ante assessment: appraisal	Peer review of proposals	- Peer and user review of scientific and societal quality - Relevance for priority themes (cf. foresight) - New actors (e.g. patient associations)	Prerogative of sponsor, often ad-hoc and in-house appraisal
Ex-post assessment: evaluation	Track-record (since 2000 returning as focus on 'excellence')	- Expert panels estimate strategic value - Methods to trace uptake, impact	Expert panels evaluate output, goal attainment

R&D policy evaluation, a definition

- Two basic types of evaluation:
  - Summative = focus on input/output relation
  - formative = focus on learning and structural change
- Methodology-based analysis and assessment
  - of appropriateness of S/T policy assumptions and targets
  - of related measures and their impacts
  - of goal attainment

Impact dimensions of public R&D spending

Main domains of impact of public spending	*Direct impacts*		*Indirect impacts*	
	Short-term	*Long-term*	*Short-term*	*Long-term*
Science (“Wissenschaft”)	e.g. scientific findings	e.g. knowledge	e.g. improved teaching	e.g. industrial spill-overs
Economy and society	e.g. improved technology	e.g. improved technical know-how	e.g. increased productivity	e.g. improved competitiveness
Policy	e.g. improved under-standing	e.g. problem-solving	e.g. problem awareness	e.g. increased general satisfaction

Source: Airaghi et al . 1999

Don’t miss the spill-overs

Sales of innovative product
Reduced process costs
Licence income

Use of technology in other parts of the business
New contacts/networks & prestige
Organisation and method learning
Competence & training

Spillovers to non-participants
User and social benefits

Figure 5. The Iceberg Model

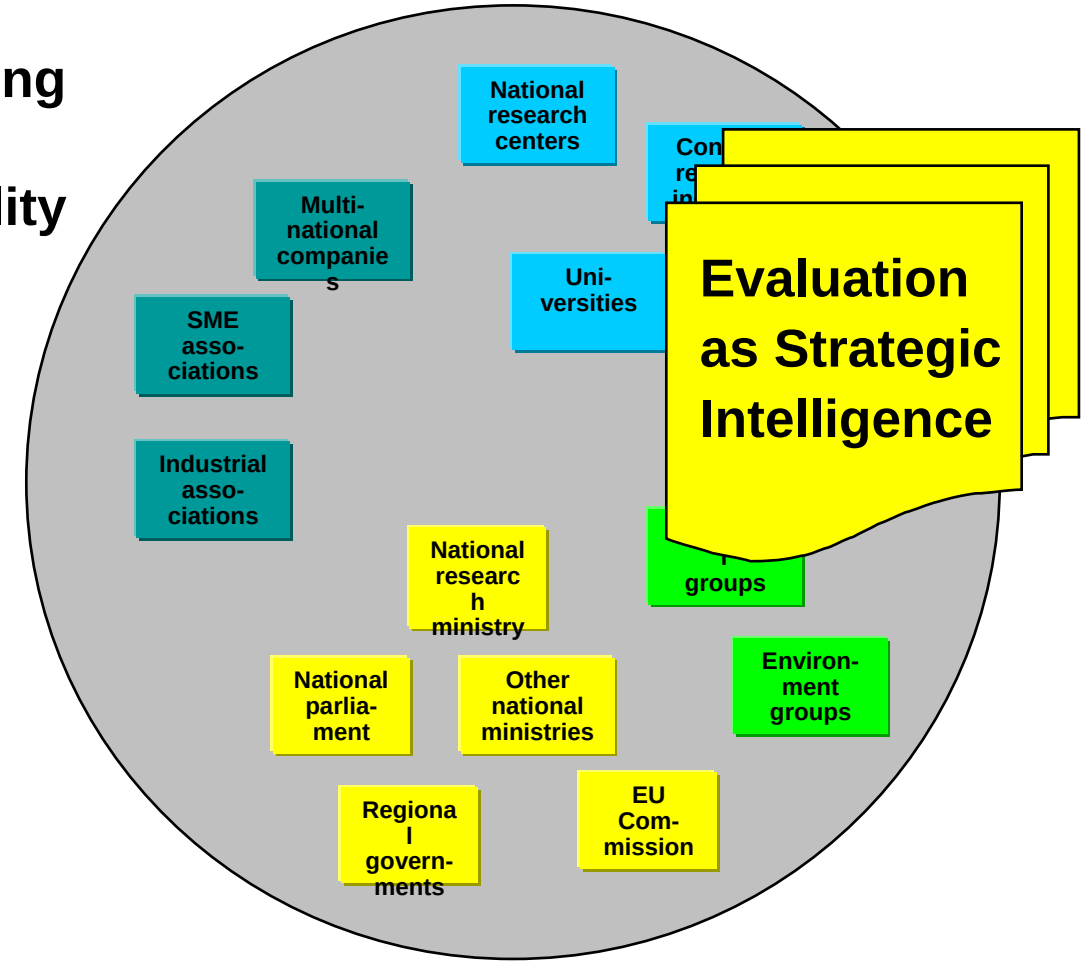
Source: Georghiou, 2007

Search for impacts of public policy

- Evaluation → identification of impact of (public) action
  - scientific, technological, economic, societal, political, ...
  - past/future, direct/indirect, intended/non-intended, ...
- Condition: Model of input/output relation, of cause/effect, actors and beneficiaries ...
- "Impact" → a rational construction of more or less complexity

R&D policy arena – a heuristic

- **Organised actors: differing interests, values, and power; bounded rationality**
- **Competition for impact and resources**
- **No dominant player?**
- **Contested policies**
- **Search for (some) alignment and policy learning - otherwise 'exit'**
- **'Enlightenment' through 'Strategic Intelligence'**



R&D evaluation methods: a “metrix”?

Evaluation mission

- ex ante (based on foresight?): strategic options?
- monitoring, real time: management, fine tuning
- ex post: learning, legitimisation
- *summative*
- *formative*

Specify mix depending on policy issue -- no general metrix!

Evaluation data

- official statistics (R&D, patents)
- bibliometrics
- questionnaire-based surveys
- interviews
- case studies

Evaluative methodology

- peer review, peer panels
- input/output; cost/benefit; before/after (descriptive statistics; econometrics)
- comparison groups
- benchmarking
- network analysis
- foresight
- technology assessment

See also "Evaluation Toolbox": ftp://ftp.jrc.es/pub/EURdoc/eur20382en.pdf

Evaluation methods

- **Quantitative: Statistical data analysis**
  - **Innovation Surveys: basic data describe the innovation process, using descriptive statistics**
  - **Benchmarking: comparisons based on a relevant set of indicators across entities**
- **Quantitative: Modelling methodologies**
  - **Macroeconomic modelling and simulation: broader socioeconomic impact of policy interventions**
  - **Microeconometric modelling: effects of policy intervention at the level of individuals or firms**
  - **Productivity analysis: impact of R&D on productivity growth at different levels data aggregation**
  - **Comparison group approach: effect on participants using statistical sophisticated techniques**
- **Qualitative and semi-quantitative methodologies**
  - **Interviews and case studies: direct observation of naturally occurring events to investigate behaviours in their indigenous social setting**
  - **Cost-benefit analysis: economic efficiency by appraising economic and social effects**
  - **Expert panels/peer review: scientific output relying on the perception of peer scientists**
  - **Network analysis: structure of cooperation relationships and consequences for individuals and their social connections into networks**

Source: RTD Evaluation Toolbox 2002

Overview

1. Purposes, options, limitations of R&D policy impact evaluation
2. **R&D policy and impact evaluation for Europe**
   - Strengths and weaknesses of FP evaluation practice
   - New perspectives: Knowledge dynamics, stakeholders, ERA
3. Outlook: Strategic evaluation

FP evaluation practice: Weaknesses

- E.g. Court of Auditors Report 09/2007:
  - Ex post evaluation too early
  - Commission data collection insufficient (late, problems with IT support)
  - Under-use of advanced methods (e.g. bibliometrics)
  - No general guideline for R&D policy evaluation (but: 'RTD Evaluation Toolbox', etc.)
  - Lack of explicit intervention logic
    - *Inevitable:* Diverging interests, targets, and logics !
    - *Instead:* Need for clear political goals, to be reached via various interventions – Need for strategic evaluation !

FP evaluation practice: Strengths

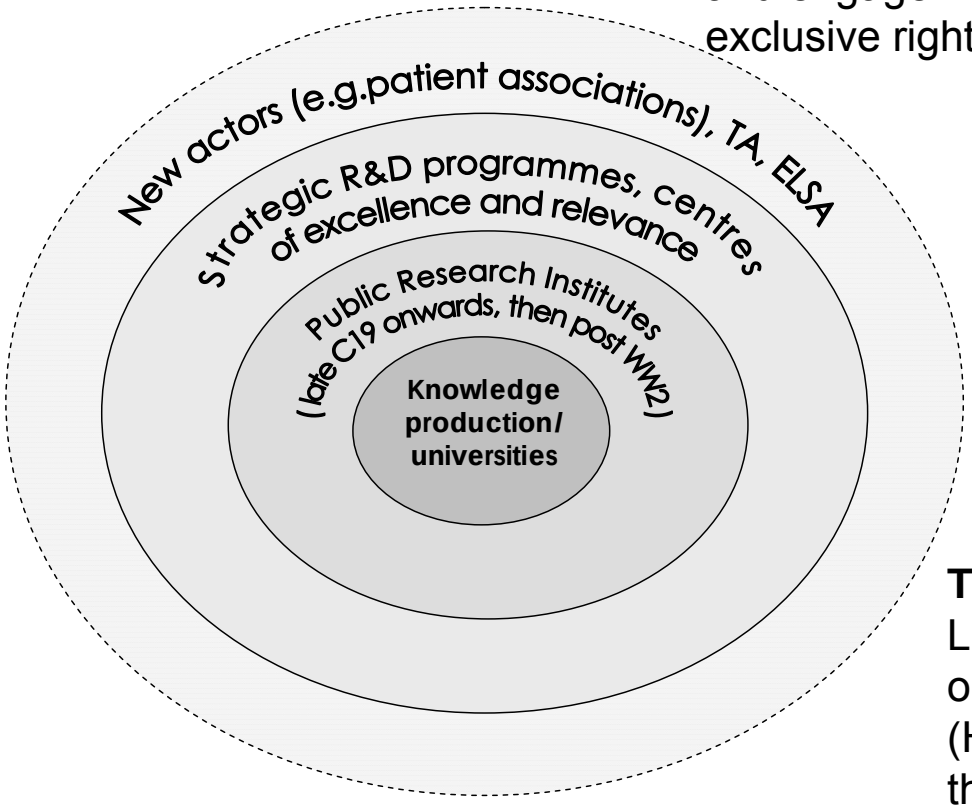
- While insufficient, EU FP evaluation practice still model case in some member state context, e.g.
  - Ex ante appraisal of FP projects
  - Some model case evaluations of FP Specific Progs
  - RTD Evaluation Network (DG Research), since 1997
  - Ex ante Impact assessment for FPs

Established eval. procedures challenged

- Research fields follow **own knowledge dynamics** and require **specific policy mixes** and **own criteria** and procedures, to do justice to their research practices:
  (e.g. medical and health research, technical sciences, humanities, social sciences; multi-, inter- and transdisciplinary research)
- Governments and administrators, from their responsibility for research output and societal goals tend to look for **simple indicators** to make (allocation) decisions.

Coping with new actors and interests

Patient associations influence research agendas and engage in research themselves, undermining the exclusive rights of scientists



Also **NGOs** bridging science and the economy, science and the community

Outreach, public engagement – feedback into research agendas?

Technology Assessment, Ethical, Legal & Social Aspects surround ongoing science and technology (Human Genome Project initiated this)

Authority over science (knowledge production) is also claimed by non-scientists (from USA Congressmen to patients and indigenous people); counter-authority is not the answer.

Assessing ERA – strategic questions (1)

- Competition across Europe – for reputation, people, funds, industrial collaboration?
  - Reduction of redundancy of research activities in Europe? Loss of diversity?
  - Crowding-out effects? Regional shifts? Winners, losers?
  - Different knowledge dynamics and clusters of excellence – changing productivity, thematic and regional distribution?
- Political and institutional responsibilities – concentration, decentralisation, shifts?
  - Pan-European standards for "excellent science and research"? Common data formats and collection procedures?
  - New governance in research institutions (funding rationales, excellence criteria, evaluation procedures, career paths)?
  - Interaction of ERA policies with national policies towards internationalisation?

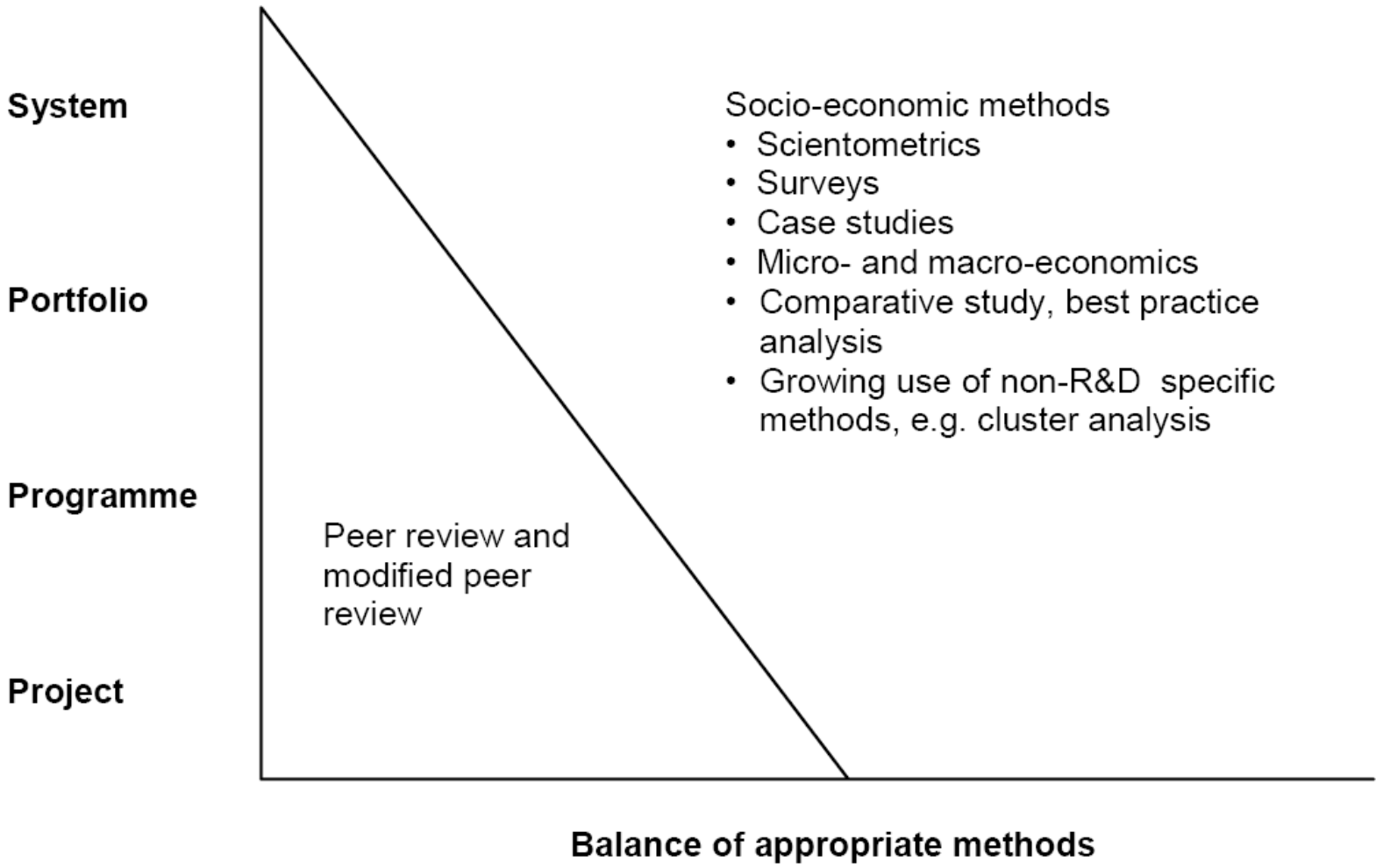
Assessing ERA – strategic questions (2)

- ERA in a global context – Contribution of policy instruments to make Europe an attractive, competitive, and connective place?
  - Usefulness of ERA (incl. national 'intra-European') policies to position Europe well in global competition?
  - Assessing ERA efforts towards 'Grand Challenges' and 'Great Ideas' (effects; governance; …)?
  - Contribution of policy instruments to mobilise vibrant links with non-OECD world?

Overview

1. Purposes, options, limitations of R&D policy impact evaluation
2. R&D policy and impact evaluation for Europe
3. **Outlook: Strategic evaluation**

Evaluation methods at different system levels

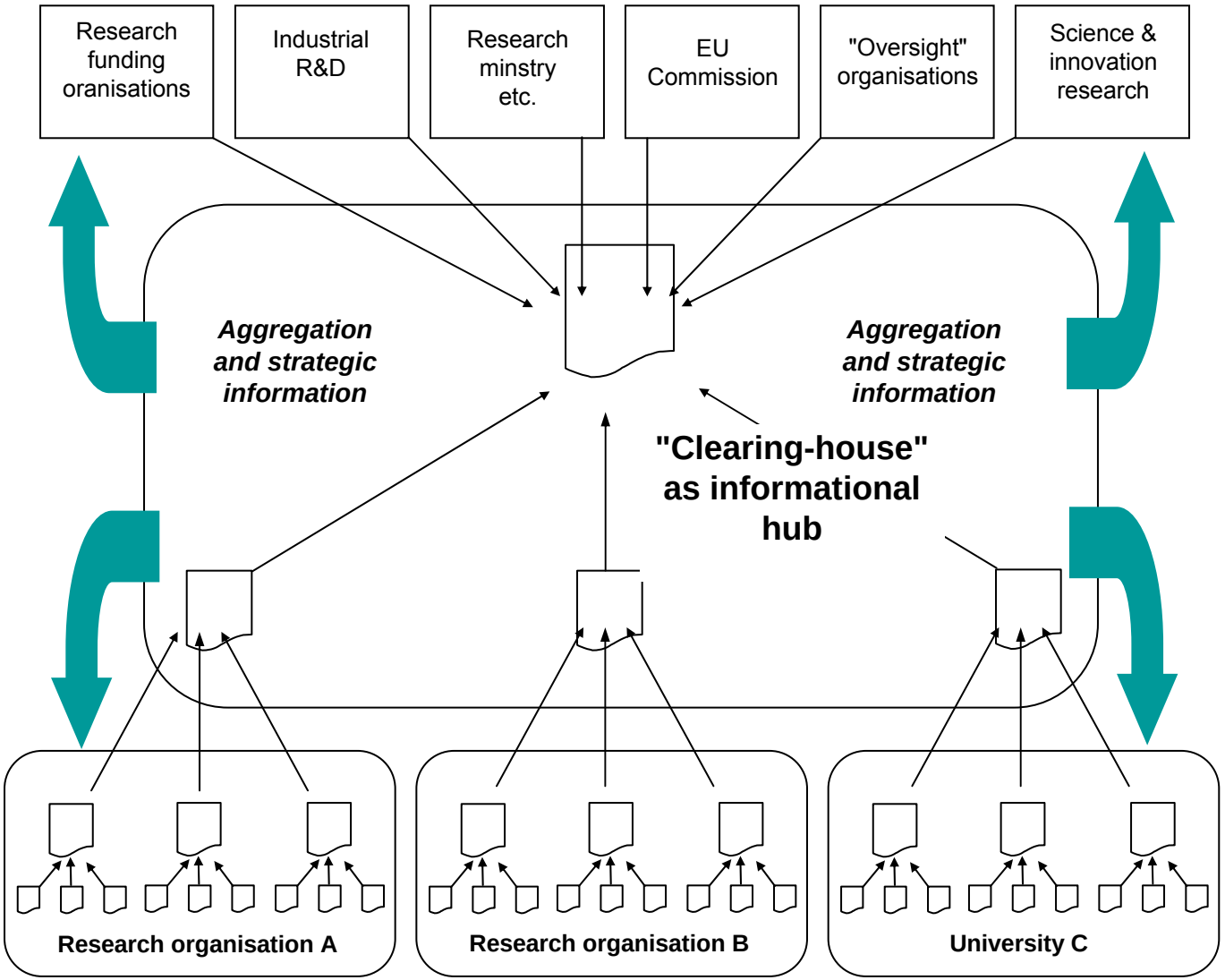


Source: Arnold 2007

Assessing ERA – strategic needs

- Need for improved strategic intelligence and indicators to understand the actual dynamics (see e.g. LEG 2008)
  - Involvement of relevant stakeholders in set-up of questions, indicators, procedures across MS and policy levels !
  - Institutional backing for data collection, treatment, analysis, and provision of evaluative information across MS and policy levels !
  - Network of "Clearinghouses" for evaluative information, warranting minimum criteria for format and quality of data, procedures of data collection and analysis, and diffusion of evaluative information

Network of "Clearinghouses"



U r s a c h e u n d W i r k u n g. - Vor der Wirkung glaubt man an andere Ursachen als nach der Wirkung.

Friedrich Nietzsche, *Die fröhliche Wissenschaft*, Aphorismus 217, 1882

Cause and impact. - Before any impact you believe in other causes than after the impact.

References

- Airaghi, A./ Busch, N.E. / Georghiou, L. / Kuhlmann, S. / Ledoux, M.J. / van Raan, A.F.J. / Viana Baptista, J. (1999): *Options and Limits for Assessing the Socio-Economic Impact of European RTD Programmes*. Report to EC, DG XII, Evaluation Unit, Brussels/Luxembourg 1999
- Arnold, E. (2004): Evaluating research and innovation policy: a systems world needs systems evaluations, *Research Evaluation*, vol. 13, 1, April 2004, 3–17
- "RTD Evaluation Toolbox" (2002): ftp://ftp.jrc.es/pub/EURdoc/eur20382en.pdf
- Georghiou, L. (2007), Avoiding the risk of under-evaluation, *Science and Public Policy*, 34(10), December 2007, pages 743–752
- Kuhlmann, S. (2003): Evaluation as a Source of "Strategic Intelligence". In: Shapira, Ph., Kuhlmann, S. (eds.): *Learning from Science and Technology Policy Evaluation: Experiences from the United States and Europe*, Cheltenham (E. Elgar), 352-379
- Leon, G. / Borras, S. / Bucar, M. / Kneucker, R. / Kuhlmann, S. / Nauwelaers, C. / Romanainen, J. (2008): *Lisbon Strategy: Between revolution and illusion - The governance challenge for knowledge policies*, Synthesis Report of the Lisbon Expert Group, Brussels (European Commission, EUR 23469)