The D4Science II experience

Andrea Manieri - ENG

ERINA+ Project

Conceptual flow of the assessment exercise

- Identify project outputs/exploitable results
- Select the most appropriate output/exploitable result
- Run the ERINA+ self-assessment
- Evaluate your impact and develop recommendations (ERINA+ experts will do this for you)

D4Science II in a nutshell

- **Main objective**

  Advance the current e-Infrastructures approach towards a Federated Knowledge Ecosystem of data Infrastructures (e.g. Genesi-DR/DEC, Inspire, FAO services, AquaMaps, Driver, OpenAire) .

More at www.d4science.eu!

Project outputs
(aka exploitable results)

- **several VREs (virtual Research Environments)** as a mean to integrate scientists and research communities in a common Virtual Environment (e.g. Aquamaps)
- **gCube,** a service-oriented middleware exploiting grids (gLite) and exposing cloud-like platform services
- **gCore,** the framework for developing services running over gCube-based infrastructures

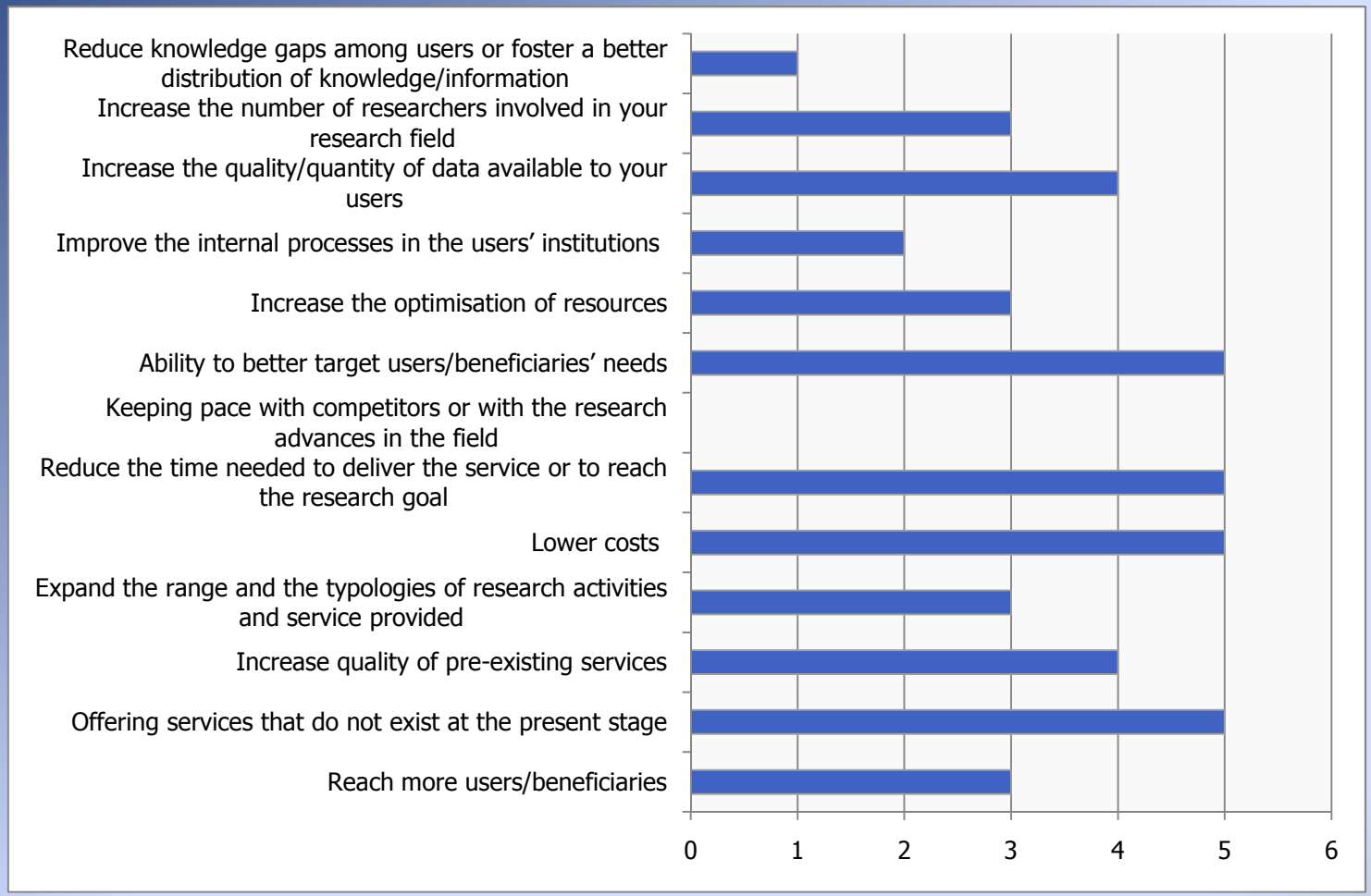
Direct beneficiaries

Expected Outputs	Direct beneficiaries	Sector specification
Virtual Research Environments (VREs)	***Researchers***	HEP scientists, biodiversity scientists, fisheries and aquaculture resources management scientists
gCube, VREs	***e-Infrastructure providers and Infrastructure providers***	Earth observation, High Energy Physics, Biodiversity and Aquaculture
gCube, gCore	***Service providers and application developers***	Digital Libraries developers and service providers; Niche Modelling developers providers of services for species distribution predictions; statistical data curation, reallocation, and publishing
VREs	***Research institutions, Commercial companies***	FAO, WorldFish Center, CERN

Indirect beneficiaries

Expected Outputs	Indirect beneficiaries	Sector specification
VREs	Research Institutes, Universities, Research Centres	Earth observation, High Energy Physics, Biodiversity and Aquaculture
VREs	Public sectors	Statistical information service (e.g. Eurostat), Maritime Affairs and Fisheries (e.g. DG-Mare)
gCube, gCore	Large companies	Software integration, PaaS, SaaS
VREs	SMEs	Scientific data harmonization and curation, Prediction of Species distribution modelling to measure the impact of climate changes and disasters

Project benefits (self-assessment)



Selected project outputs

- **the AquaMaps VRE**

  A unique virtual environment providing set of services for the generation, standardized dissemination, and mapped visualization of model-based, large-scale predictions of currently known natural occurrence of marine species.

Areas	**Describe the output**	**At the end of the project**	**Within the next 3 years**	**Within the next 5 years**
Number of new products	1. Species distribution maps 2. Biodiversity maps 3. Species distribution maps raster x GIS 4. Biodiversity maps raster x GIS 5. transect	6. map comparison 7. long term forecast 8. extension to fresh water species	More prediction algorithms to be applied	

Identify an ex-ante scenario

- Current process for data collection, analysis and map preparation
  - Manually collected data
  - Intensive manual labour in data analysis
  - Compute demanding algorithms for data processing and map generation

Impact on direct beneficiaries

- Labour/Time savings for the elaboration of one dataset
  - ex-ante: 6 temporal units (e.g. 2 persons * 3 months, 6 persons for 1 month, etc)
  - ex-post: 0.25 temporal unit

  Estimated savings: 99,2%

- How many datasets? How many times in 1 year?

Process Improvement

- In a given time frame (3 months) from 5k maps up to 400k
- Experimental data:
  - in 6 months, from 10k to 90k complete dataset elaboration (achieved only in 9 activities of 12 hours each). Savings: >53 temporal unit
  - Error frequency: from 1/100 to 1/90000

Cost Efficiency (running the service)

- Ex ante: 2 persons + 15% sys admin + 1 server (dedicated)
- Ex post: 1 person + 0.5% sys admin) + 20% of 10 on-demand virtual servers
- Savings: 73%

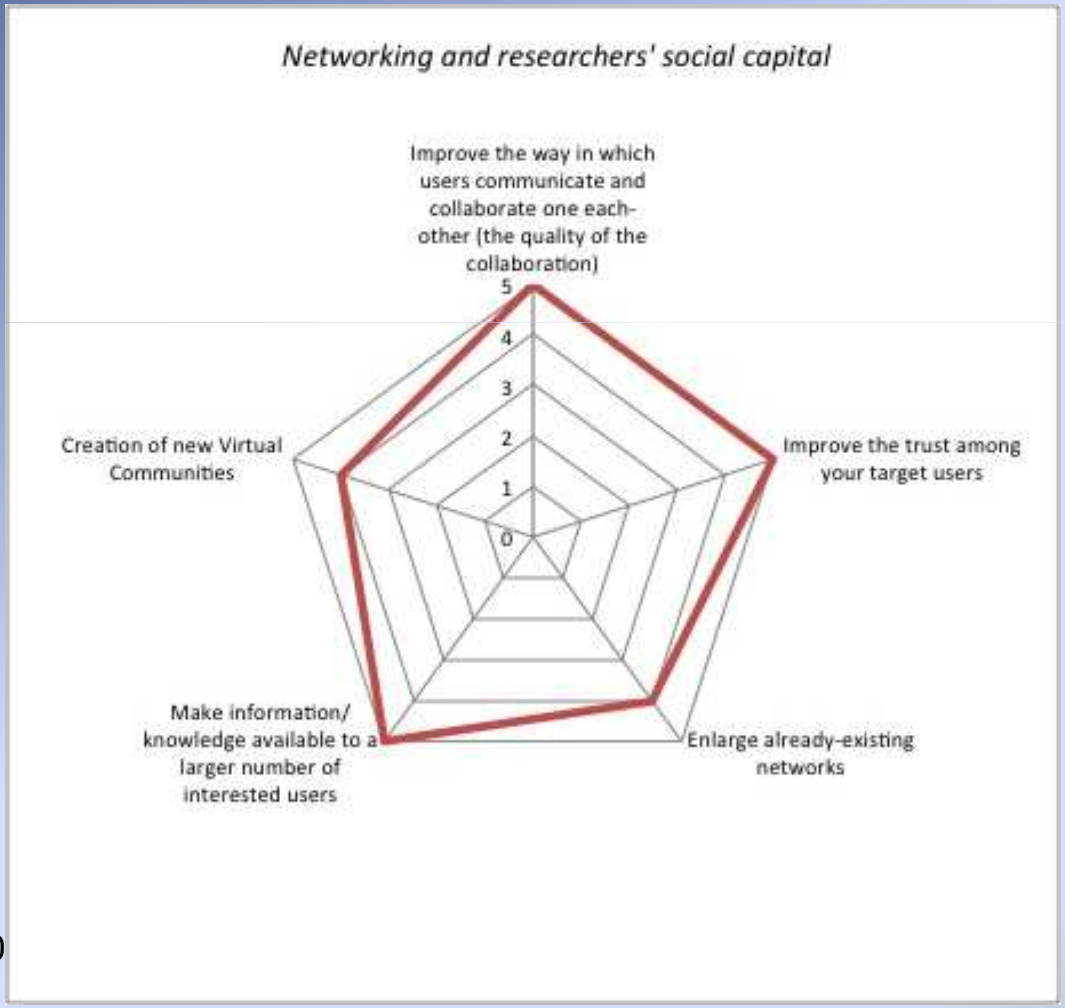
Social impact

The Social impact analysis of the D4Sceince II project presented in the following slides refers mainly to the data collected through ERINA+ preliminary questionnaire.

More data and a more complete analysis going to be performed by means of the ERINA+ second questionnaire.

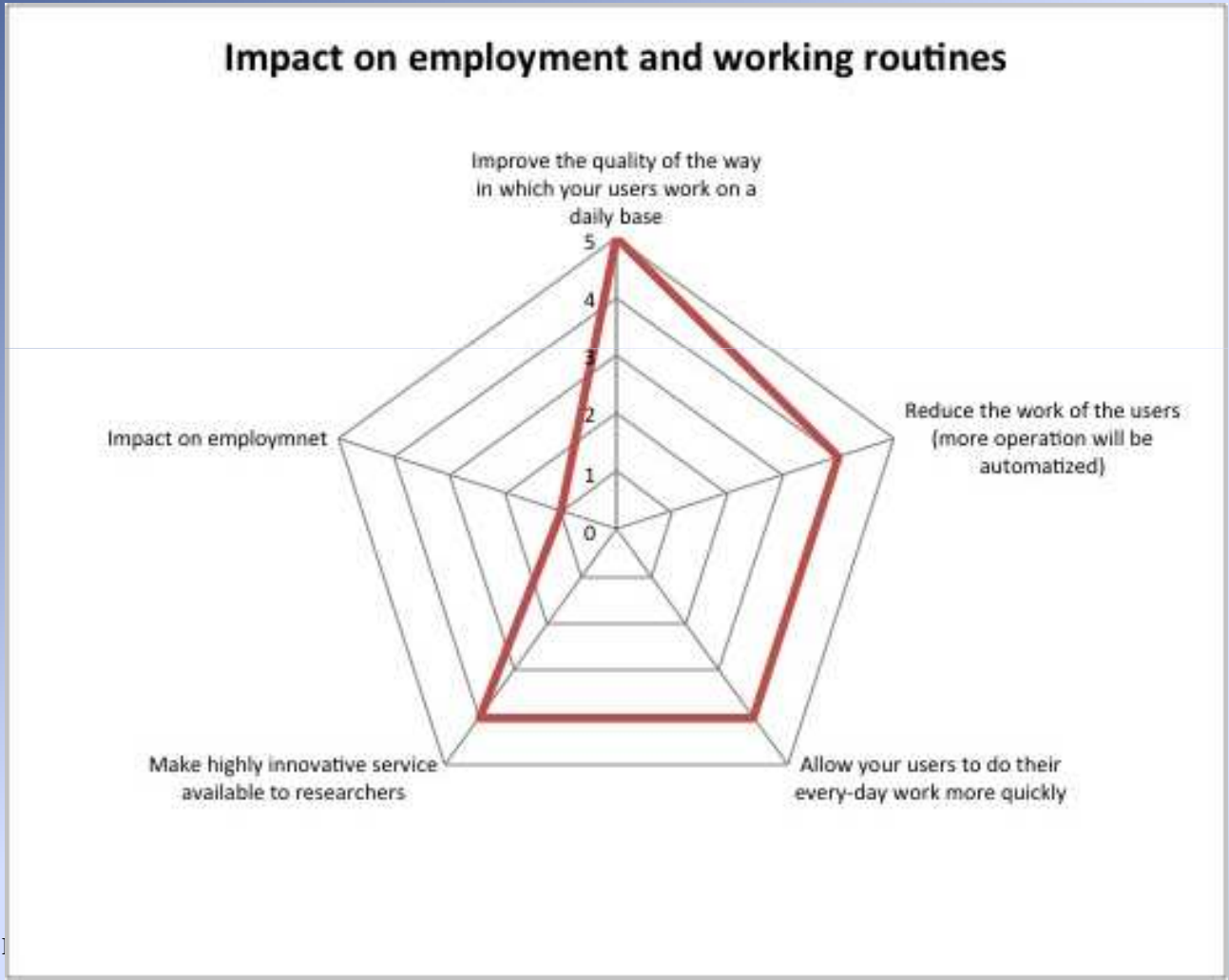
Social impact

- **10 new collaboration agreements with research institutions**
- **50 networking events**
- **10 new virtual communities**

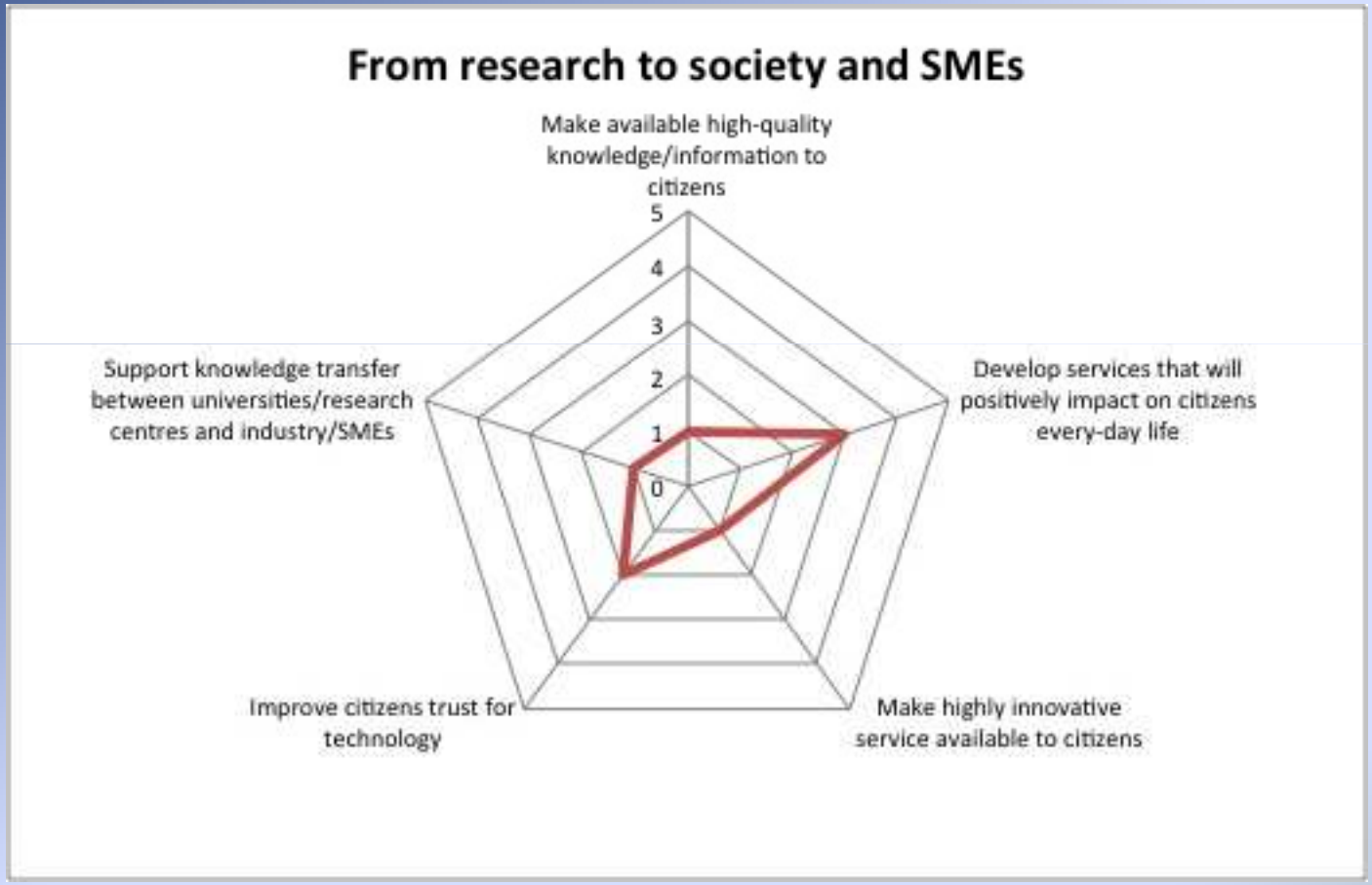


Social Impact

"D4Science-II support the creation of several Virtual Research Environments (VREs) that offer significantly enhanced services to their users without occurring in high development and maintenance costs"



Social impact



Data collection

- D4science side:
  - One day to understand the process/questionnaries and select the right people (also including project plan and personal constraints)
  - 1&1/2 day – 2 two people (Leonardo Candela, Lino Pagano – CNR-ISTI) to reply to the whole questionnaire for Acquamaps only
  - Need to go for gCube/gCore; each of VREs.
- ERINA+ side:
  - Half day – 2 people and three calls to support and explain
  - Half day – 4 people to analyse data and provide feedback

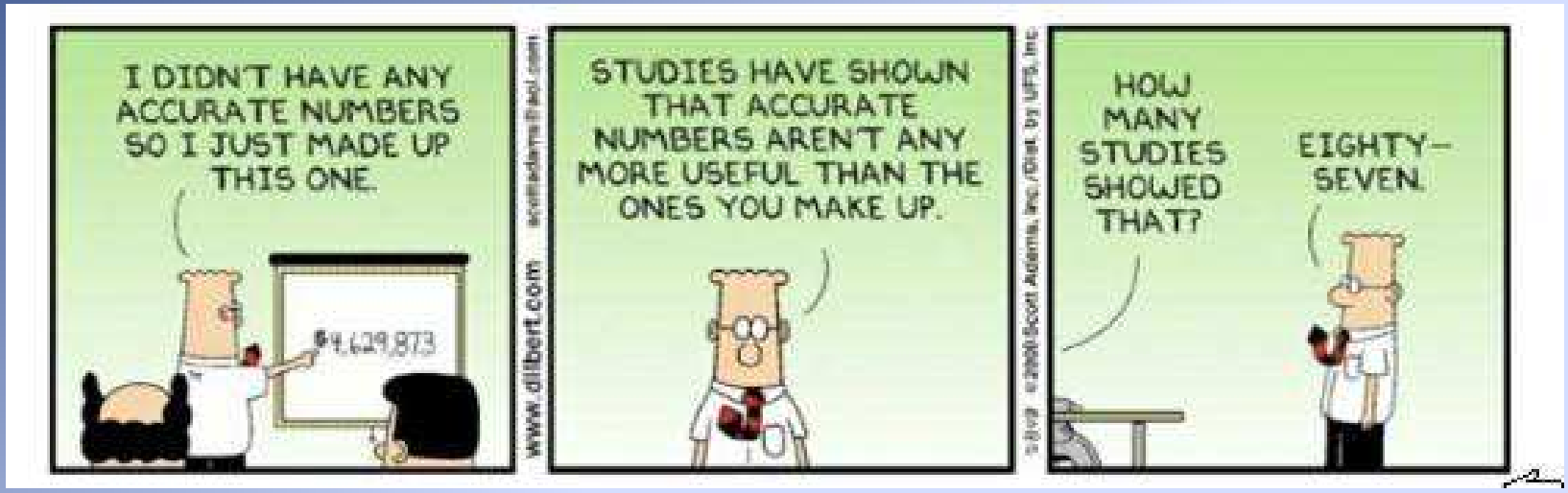
D4Science II statement

"*We appreciated the ERINA+ approach because there is a need to have a standardised system for the collection of projects' performance indicators.*

The complexity of the metric is not a problem if the methodology is agreed from the beginning of the project."

Pasquale Pagano, D4Science-II

THANK YOU


I DIDN'T HAVE ANY ACCURATE NUMBERS SO I JUST MADE UP THIS ONE.
$4,629,873
www.dilbert.com
STUDIES HAVE SHOWN THAT ACCURATE NUMBERS AREN'T ANY MORE USEFUL THAN THE ONES YOU MAKE UP.
HOW MANY STUDIES SHOWED THAT?
EIGHTY-SEVEN.