SA National Integrated Cyber-Infrastructure System ……. and eResearch

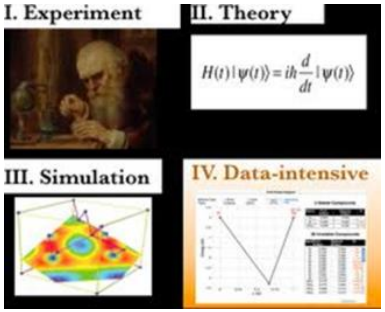
@ UCT
20141125

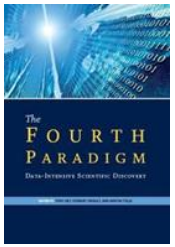
Professor Colin Wright
National Integrated Cyber-Infrastructure System
DST & CSIR Meraka Institute
South Africa
cjwright@csir.co.za

science & technology
Department:
Science and Technology
REPUBLIC OF SOUTH AFRICA

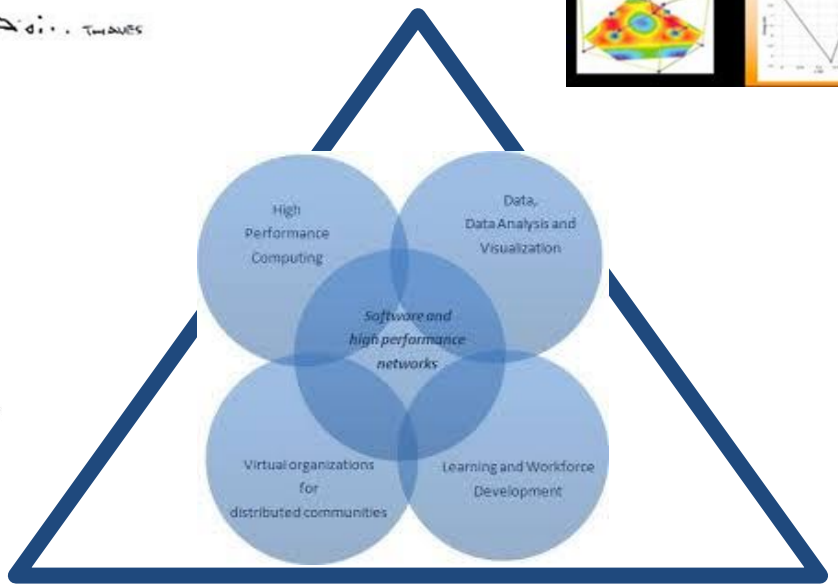
CSIR meraka INSTITUTE
our future through science
African Advanced Institute for Information & Communications Technology

The Knowledge Triangle at Work



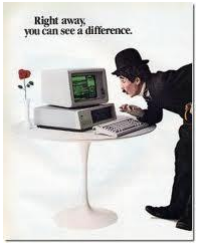


Research



Education

Innovation

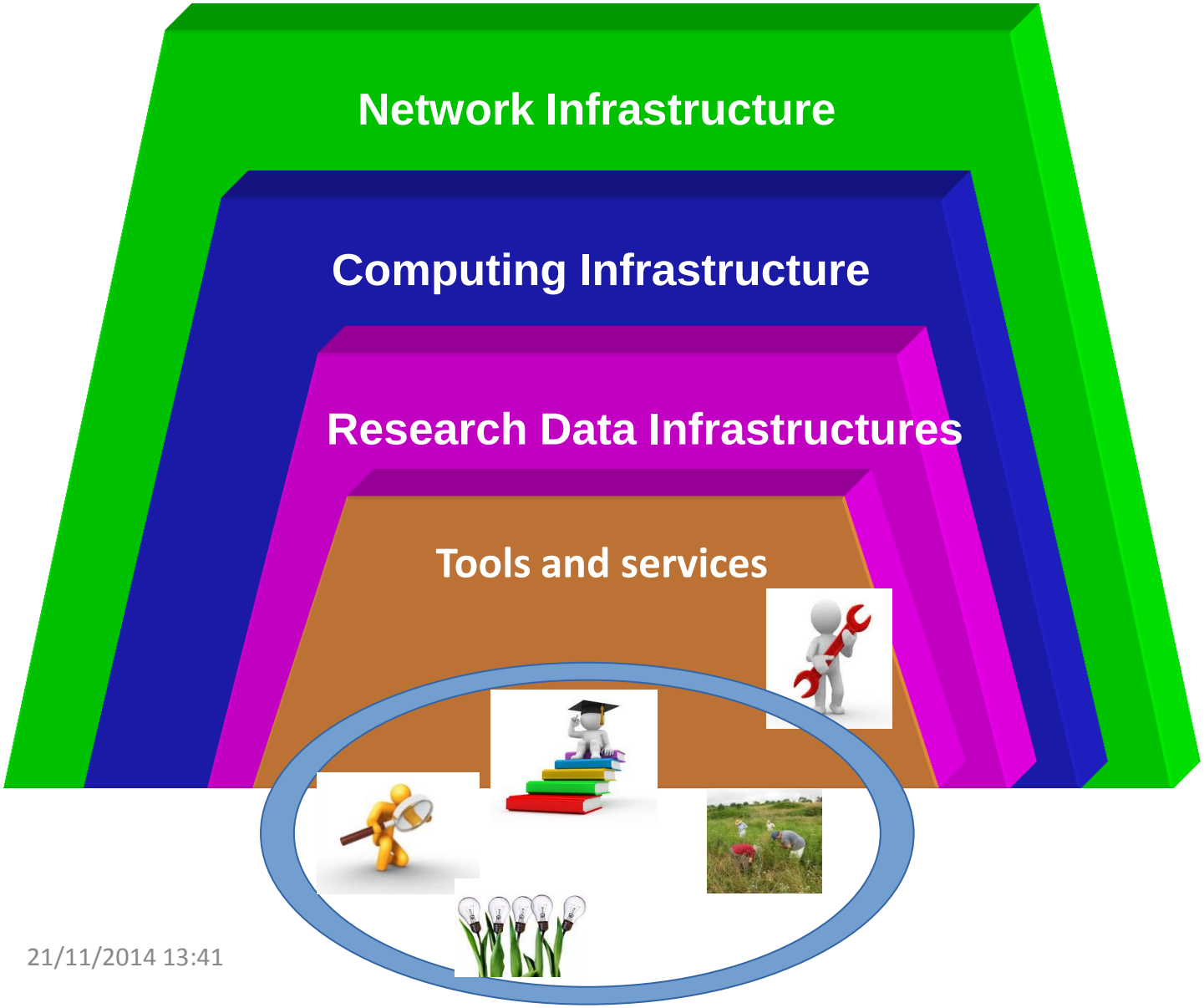


SARIR

To be genuinely competitive in the knowledge economy, one must be competitive at:

- producing knowledge through research
- diffusing it through education
- applying it through innovation

CyberInfrastructure: Core components and Activities



Support Policy

(International) Cooperation and Links

Cyberinfrastructure Ecosystem

Organizations:
Universities;
Research labs;
Museums;
Libraries;
Communities

Expertise
Research &
Scholarship;
Education;
Skills and
Training;
eScience

Research
Infrastructures
Large facilities;
satellites; Telescopes;
Colliders; Distributed
sensor arrays; Ocean
buoys;...

Computational
Resources
Supercompurs,
clusters; Grids;
Clouds; Data
Centers;
Visualisation

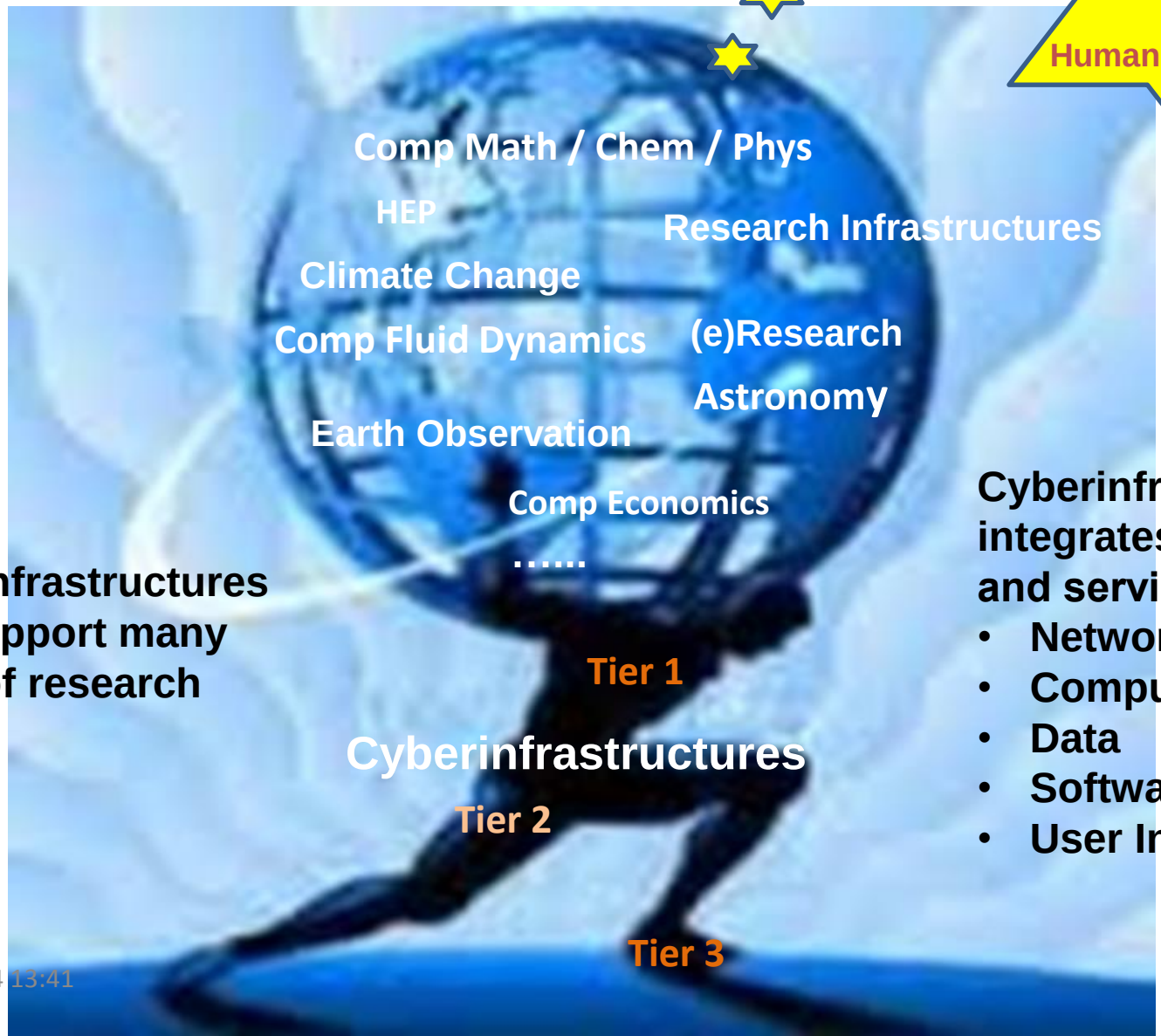
Educate
Research
Innovate

Data
Data repositories;
Collections, libraries;
Journal collections; Mining,
analysis; Visualisation;
National licences

Software
Software
development
and support;
Middleware;
Applications;
Advanced
services

Networking
Global, national, campus;
Lightpaths, GOLE; SDB;
Advanced Services

Cyber-Infrastructures domains of impact ….



Cyberinfrastructures now support many areas of research

Cyberinfrastructure integrates resources and services …..

- **Networking**
- **Computing**
- **Data**
- **Software**
- **User Interfaces**

Wikipedia

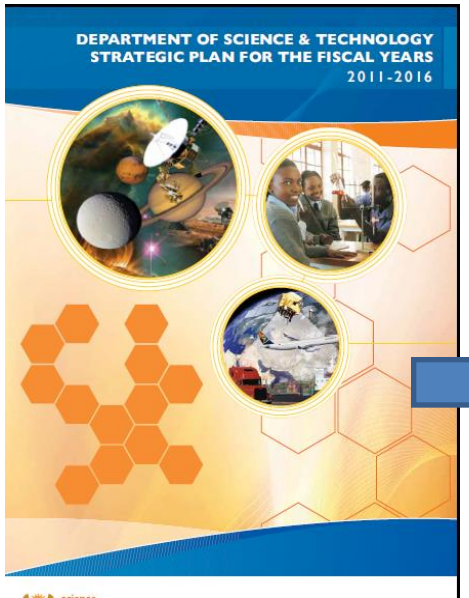
“E-Science is *computationally intensive science* carried out in highly distributed network environments, or science that uses immense data sets that require grid-computing; sometimes includes technologies that enable distributed collaboration, such as the Access Grid. (1999)

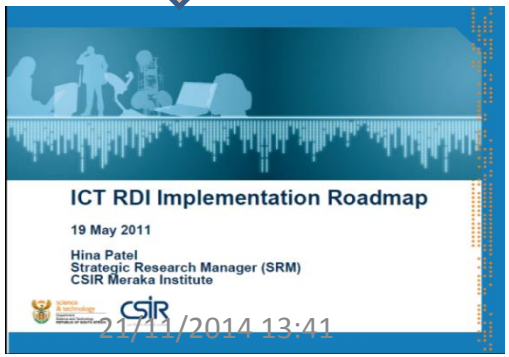
More broadly "*the application of computer technology to the undertaking of modern scientific investigation*, including the preparation, experimentation, data collection, results dissemination, and long-term storage and accessibility of all materials generated through the scientific process. These may include data modelling and analysis, electronic/digitized laboratory notebooks, raw and fitted data sets, manuscript production and draft versions, pre-prints, and print and/or electronic publications.“”

SA National Cyberinfrastructure System

science & technology
Department:
Science and Technology
REPUBLIC OF SOUTH AFRICA

Science Imperatives led DST into the CI arena





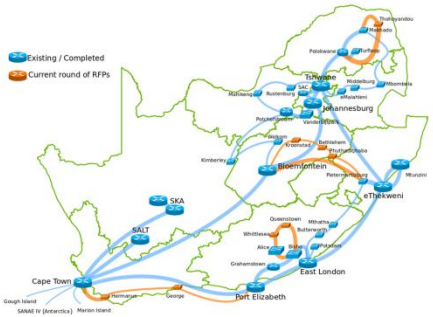
CSIR
our future through science

meraka INSTITUTE
African Advanced Institute for Information & Communications Technology

CHPC
CENTRE FOR HIGH PERFORMANCE COMPUTING

SANReN
South African National Research Network

SOUTH AFRICAN NATIONAL GRID

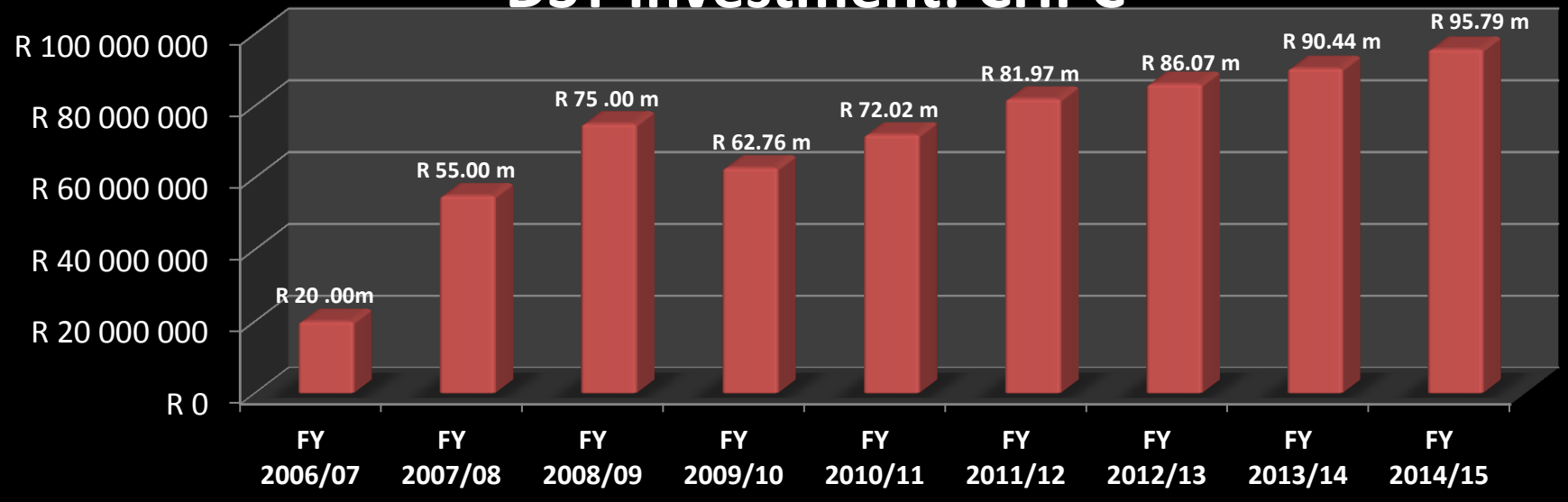


VLDB / DIRISA

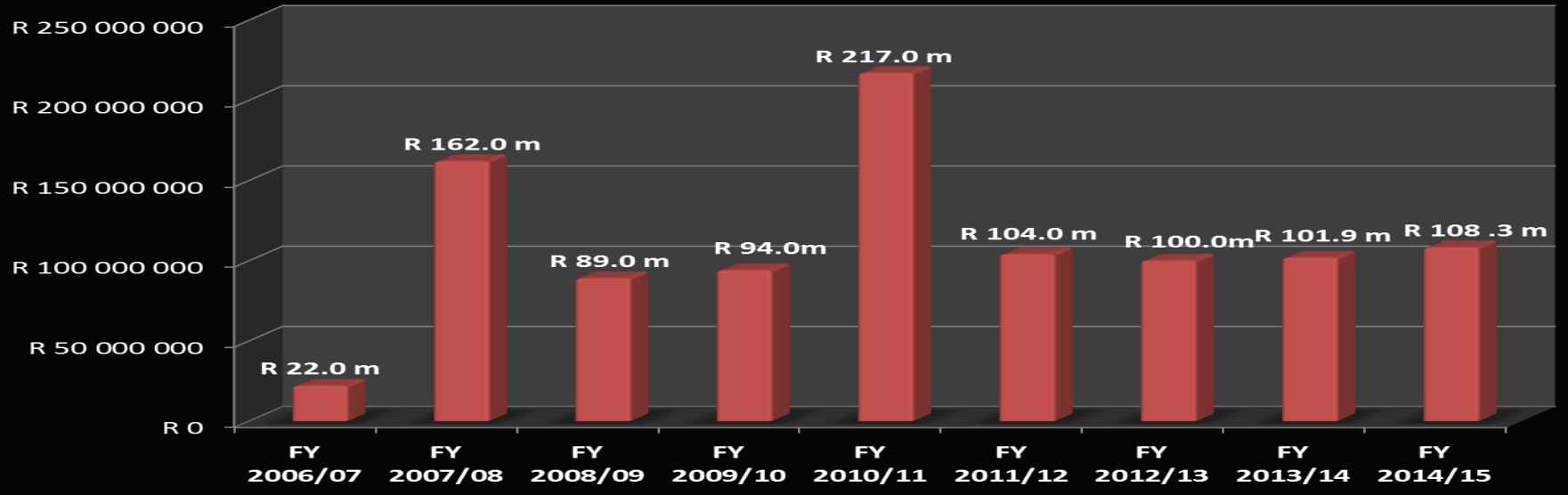
dirisa

CSIR
our future through science

DST Investment: CHPC



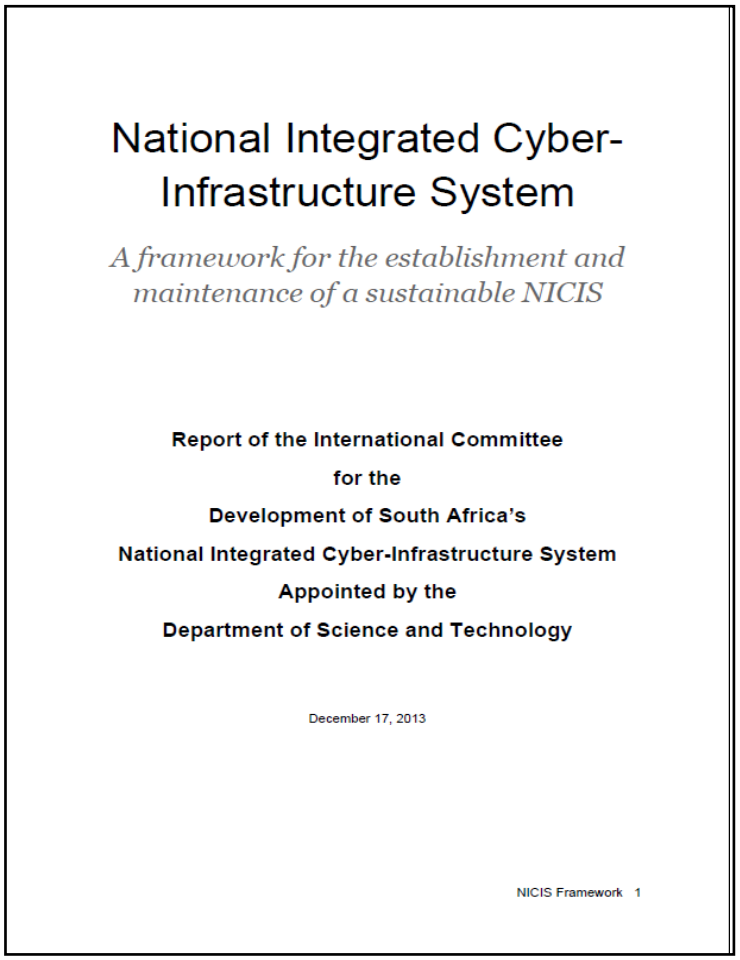
DST Investment: SANReN



+ WACS, VLDB / DIRISA. SAGrid no formal DST support as yet.

2012—2014 NICIS Review

CI, together with RIs, are fundamental pillars in developing a competitive Research, Education and Innovation system



National Integrated Cyber-Infrastructure System

A framework for the establishment and maintenance of a sustainable NICIS

Report of the International Committee
for the
Development of South Africa's
National Integrated Cyber-Infrastructure System
Appointed by the
Department of Science and Technology

December 17, 2013

NICIS Framework 1

Timeline of events:

- Review Committee appointed by Minister: July 2012
- NICIS Review Committee submit Report: December 2013
- CI Specialist appointed to advise DST: April 2014
- NICIS report made public mid 2014
- DST EXCO considered recommendations and agreed on responses: June 2014
- Consultation with Minister of S&T: agreed
- Presently: analysis and deliberations underway regarding the location, structure and governance of NICIS
- Launch new NICIS 20150401

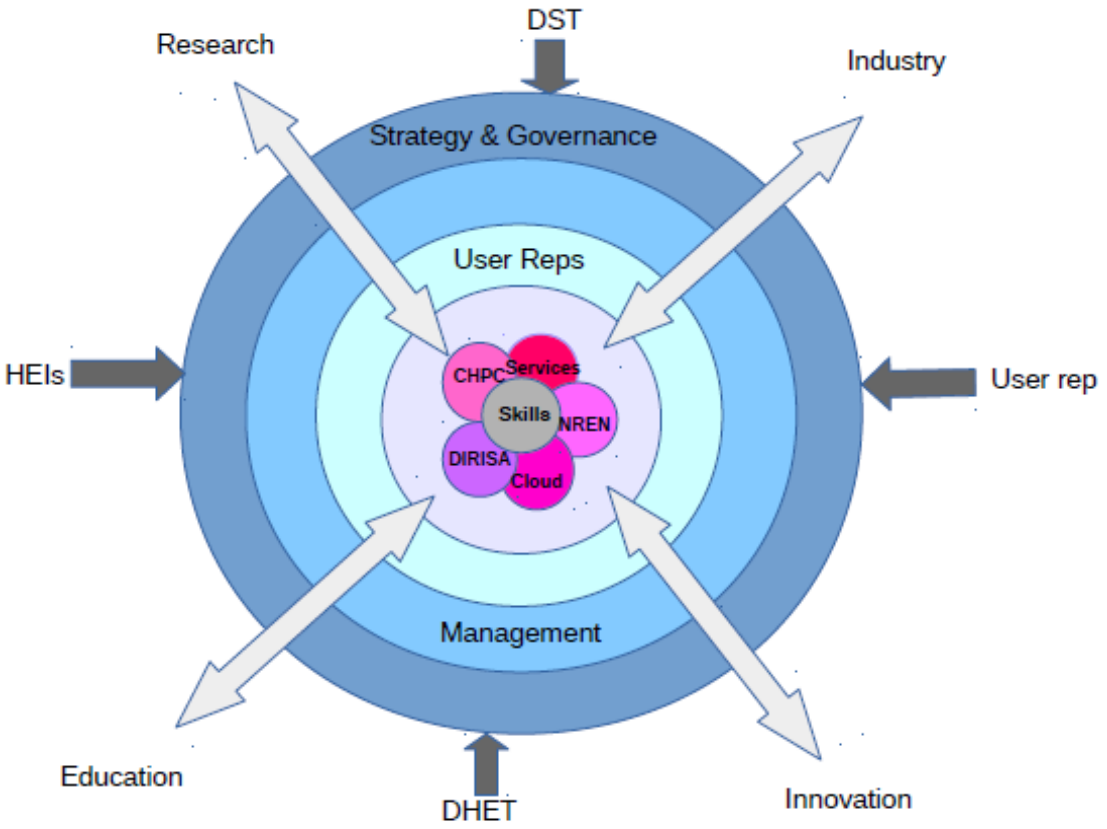
NICIS
National Integrated Cyber-Infrastructure System

NICIS VISION: take national leadership in the provision of a comprehensive Cyber-Infrastructure essential to 21st century advances for South Africa in research, education and innovation.

NICIS MISSION: increase knowledge creation through provision of a national platform of essential Cyber-Infrastructure.

NICIS PRINCIPLES

- Joint planning and budgeting
- Good governance
- Visibility of CI services
- Sustainability
- Constructive stakeholder engagement



TIER 1

CHPC CENTRE FOR HIGH PERFORMANCE COMPUTING

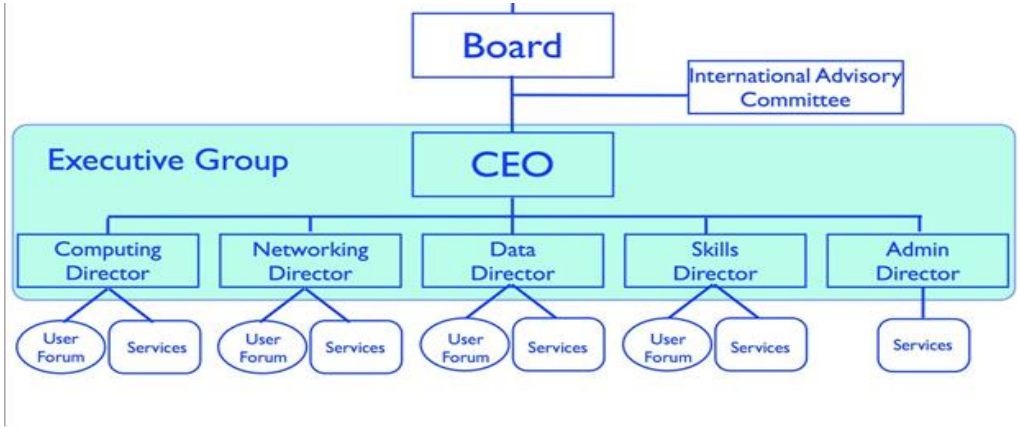
SANReN South African National Research Network

dirisa

Skills & Training

SOUTH AFRICAN NATIONAL GRID

Current discussion and negotiation underway re NICIS location, legal structure, governance and management



- A national “Cyber-Infrastructure” platform
- Provide raw bandwidth, compute cycles and data storage and curation; through to value-added services; cloud and grid services; advanced user support and training activities;
- Skills and Training Service Area coordinate cross-cutting HCD theme interactively with HEIs;
    - NICIS provides glue to ensure coherent approach to CI training and education.

Centre for High Performance Computing

Tier 1 level collaboration with Tier 2 computing ecosystem

NICIS SC: CHPC *in essentially its current form should take on the role of the Computing Services area, with some changes to its mandate. More formal processes to assess user satisfaction and needs must be established with requisite metrics for performance.*

- **Compute-intensive**,
- **Communication-intensive**
- **Data-intensive**
- **Sustainable, impactful, user sensitive**
- **Visualisation, Cybersecurity**

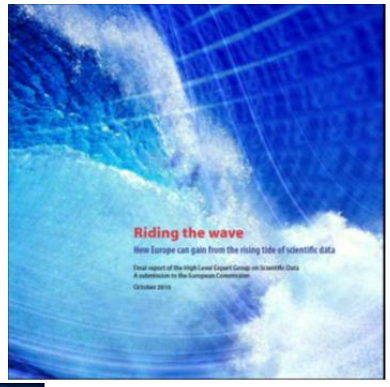
NICIS SC: Networking Services: NREN

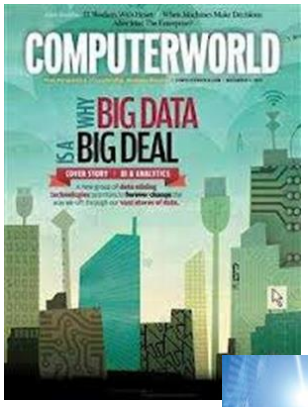
- *Important for DST to have significant involvement in the NREN both because it will be a primary funder and so will need to account for the spending of public funds and because it is in the position of having strategic oversight and can take into account all national interests.*
- *Other key parties such as the universities will also have a stake in the process and be able to hold the service area to account.*

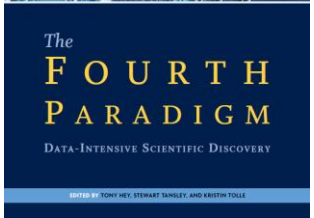
DST response: SANReN South African National Research Network + TENET Tertiary Education and Research Network of South Africa

- NREN = SANReN + TENET
- Maintain current division of responsibility and collaboration between SANReN and TENET; and
- Re-affirm / establish a comprehensive MoU for sustainable regulation of SANReN-TENET relationship.

And then there was DATA!! Data is the new oil … !







RDA
RESEARCH DATA ALLIANCE

www.csir.co.za

Some history: 2008 VLDB Proposal summary

Establish a South African data curation, warehousing and analysis facility, staffed by people with data and software skills to support discipline specific researchers. Structure as a national facility with accessibility and availability to all SA Researchers. US cyberinfrastructure report strong point that whilst effective and successful curation is dependent on hardware, software and connectivity, it is even more dependent on skilled and quality people.

Objectives

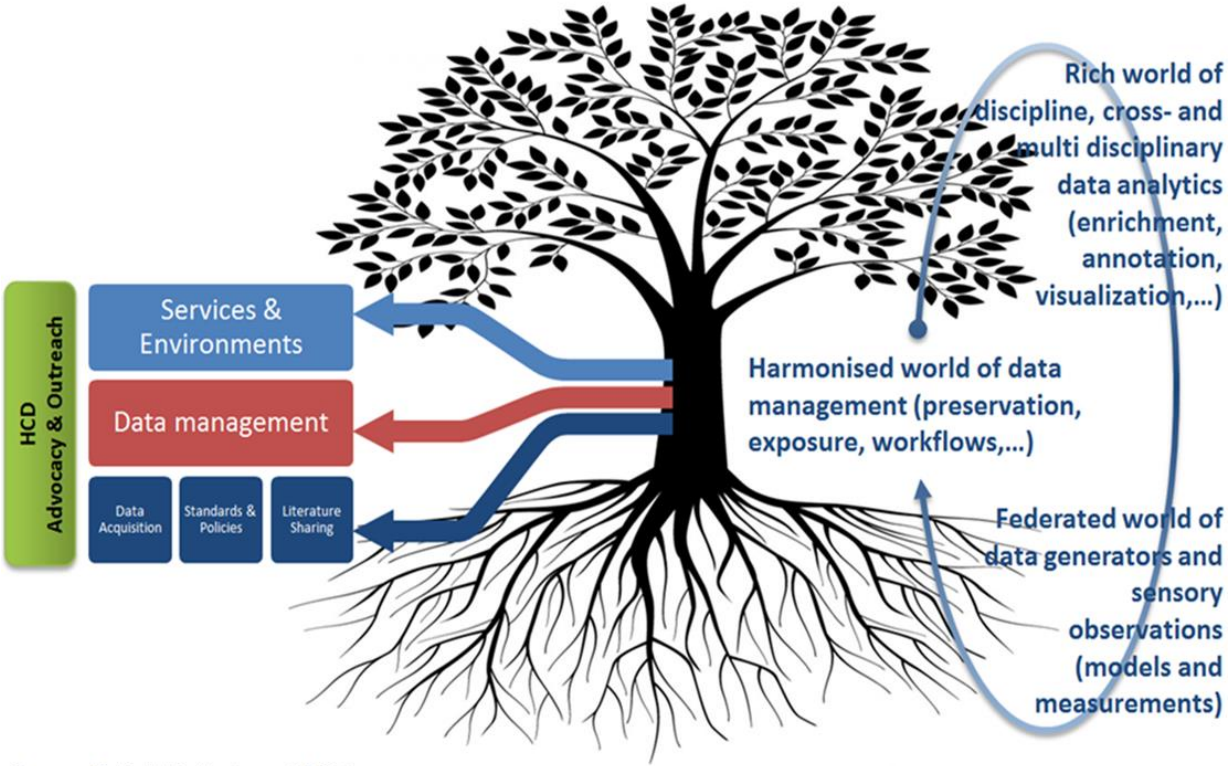
- Effective governance;
- S&T excellence;
- Human Resource Development;
- Innovation and Learning:;
- Funding sustainability;
- National Facility;
- Partnerships: Customers and Stakeholders
- Government support;
- Operational excellence: world-class project management and research processes.

Co-located with CHPC

NICIS SC: Data Services

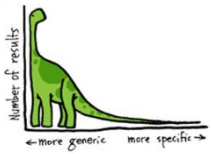
- ***NICIS Data Services Area (should) be leading organisation within South Africa to advocate for and implement data initiatives across the research community.***
- *NICIS (should) work with the community to develop an ambitious proposal on data services to DST.*
- *New funding necessary for NICIS…expanded Data Services.*

- Recommendations fully endorsed – DIRISA to be developed into a full NICIS entity;
- Strategic approach to Treasury around Big Data funding.

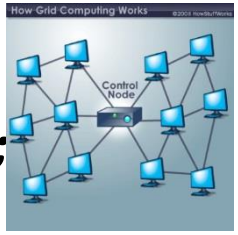


Source: Colin Wright, August 2014

- Expansion in data-intense portfolio requires infrastructural support;
- Data analytics is becoming a key driver in E-based commercial activities.



NICIS SC : Cloud and Grid Computing



a) cloud computing to be a key function of Computing Services; SAGrid continued with modest additional support to increase robustness, DST consult with stakeholders to determine location

Cloud and Grid cross-cutting tools providing services across all the services areas. Necessary to enable NICIS to support big science, big data and e-Research.

NICIS SC: Tier 1 and Tier 2 Infrastructures

- ***… continue to focus investments in Tier 1 centralised facilities,*** *with expectation of collaboration with regional Tier 2 facilities.*
- ***… important that researchers have access to a range of HPC facilities,*** *including both national Tier 1 resources and institutional Tier 2 resources.*
- *[Other funding agencies] should fund Tier 2 …*
- *Universities to be asked to respond to report and draw up individual and collective strategies for provision of Tier 2 (and 3) facilities.*

NICIS SC: Research

NICIS should not have an extensive in-house research role. However, CI-related research should be strongly supported within the research communities outside the scope of NICIS.

National Research Foundation

higher education & training
Department: Higher Education and Training
REPUBLIC OF SOUTH AFRICA

NICIS SC: Consultation, Stakeholders and Accountability

High level of accountability to the stakeholder. Universities should be asked to respond to report. Collaboration and coordination between DHET and DST essential in prioritising mid-range facilities, connection to national leadership-class compute capabilities, human resource use, and benefit of procurement.

- Recommendations supported, improvements needed;
- Stakeholder community will be consulted as soon as current deliberations concluded.
- International best practice regarding CI governance and management indicates that stakeholders should be seriously involved and consulted in CI deliberations, and that they participate in governance structures.

NICIS SC: Skills and Training

NICIS to offer effective coordination of CI Skills and Training services within a sustainable framework. Additional funding necessary for new NICIS Skills and Training Services area.

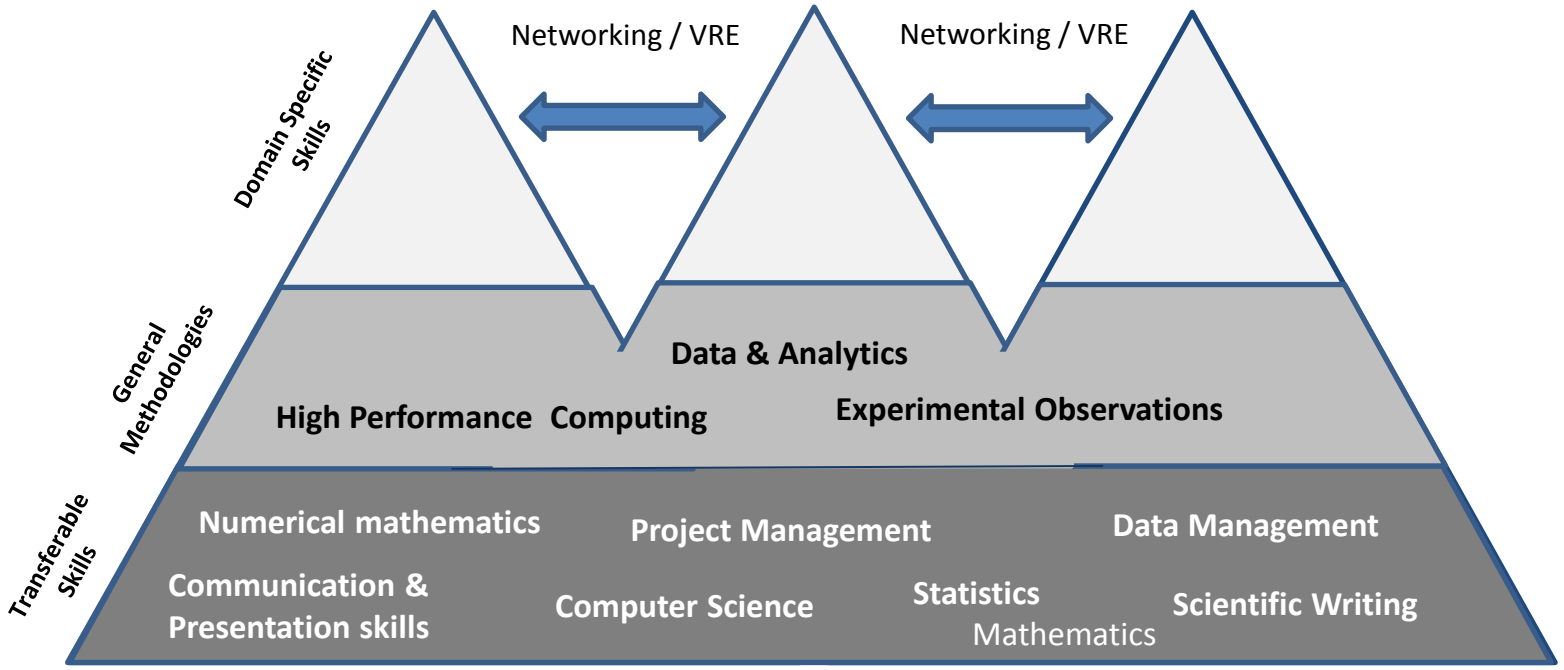
Recommendations endorsed:

- Explore piloting a multi-university / multi-disciplinary master's degree in e-science to grow number of PhDs in this area;
- Consult with HESA to ensure alignment and pipeline at undergraduate level; and
- Recognise that the impact of a Cyber-Infrastructure is only fulfilled if trained professionals are available.

Examples

- eResearch
- Data Science
- Computational ….
- HPC
- Data analytics
- …….

Skills Sets and Capabilities needed by Researchers in Academia and Industry

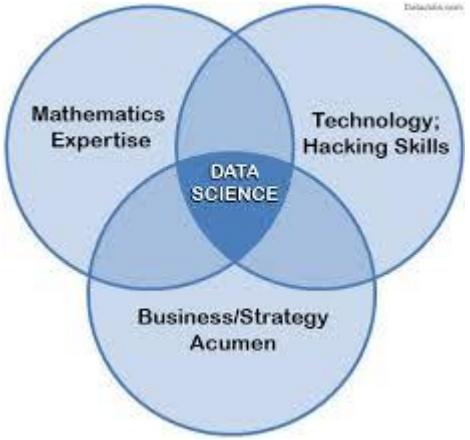


Research Communities

eScience / e-Research

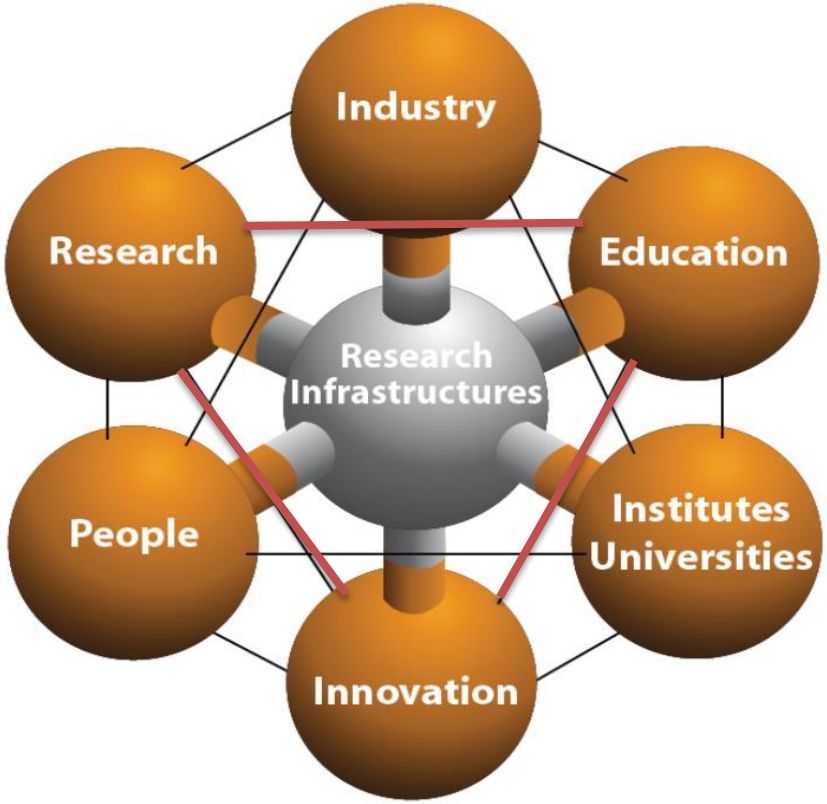
Cyberinfrastructures skills and human resources: specific aspects

1. Co-develop and coordinate courses with stakeholders; support community in developing educational activities, programs;
2. ***Effective usage of cyberinfrastructures requires:***
   - ***Mathematics, Numerical Analysis***
   - ***Computation; Simulation; Data Analytics skills***
   - ***Computer Science; Visualisation***
   - ***Software skills: software engineering***
   - ***Statistics, Big Data Skills***
3. ***Changes in research institutions***
   - ***Need for institutional leadership***
   - ***Need to support professional development***
4. ***New careers***
   - ***Professional recognition***
   - ***Both specialists and domain aware researchers***



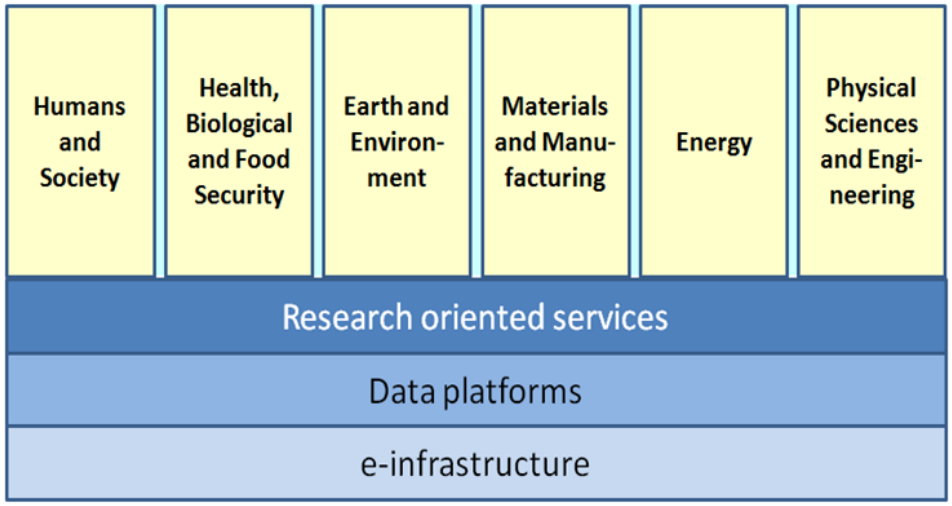
Research Infrastructures / CyberInfrastructures – a driver for Knowledge and Innovation both nationally and internationally

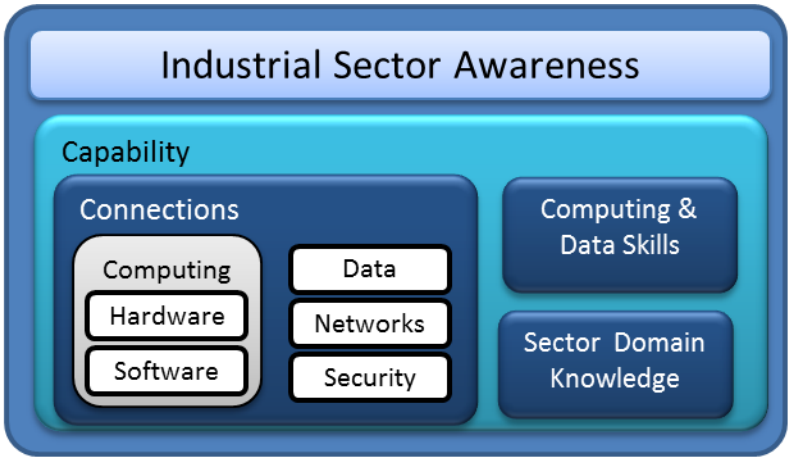
➡ However it needs knowledgeable, trained, skilled, open, committed people to make any impact.



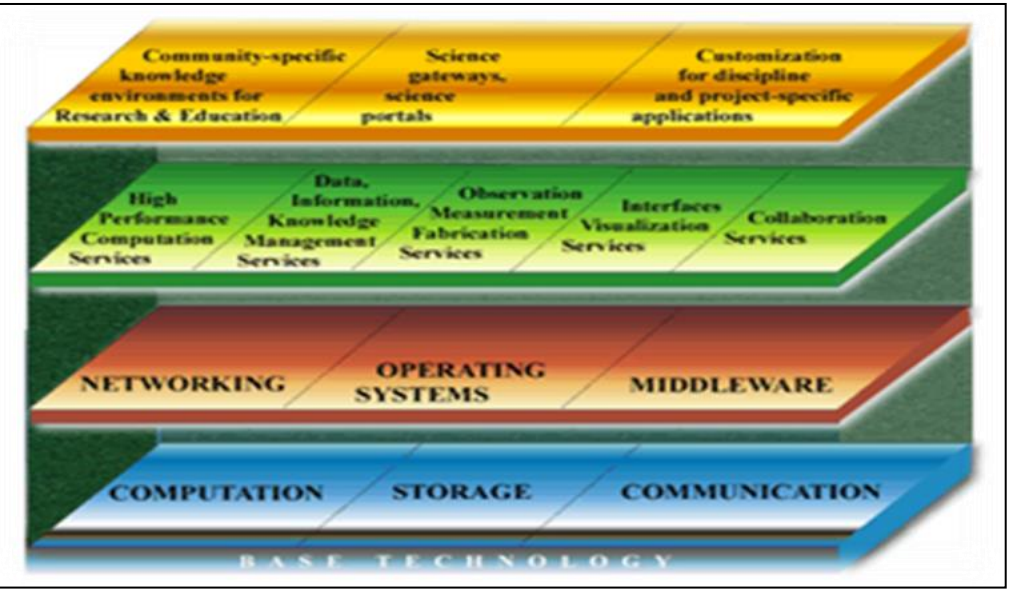
In higher education and training, South Africa ranked last in the quality of maths and science education. The country was 140th out of 144 for the quality of its education system.

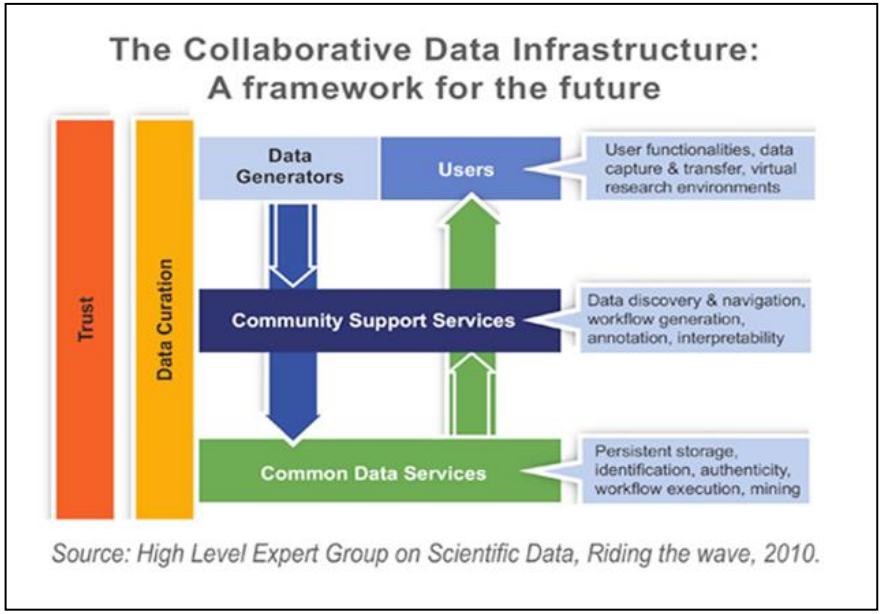
Different Perspectives





D. Tildesley: Vision of integrated e-infrastructure ecosystem

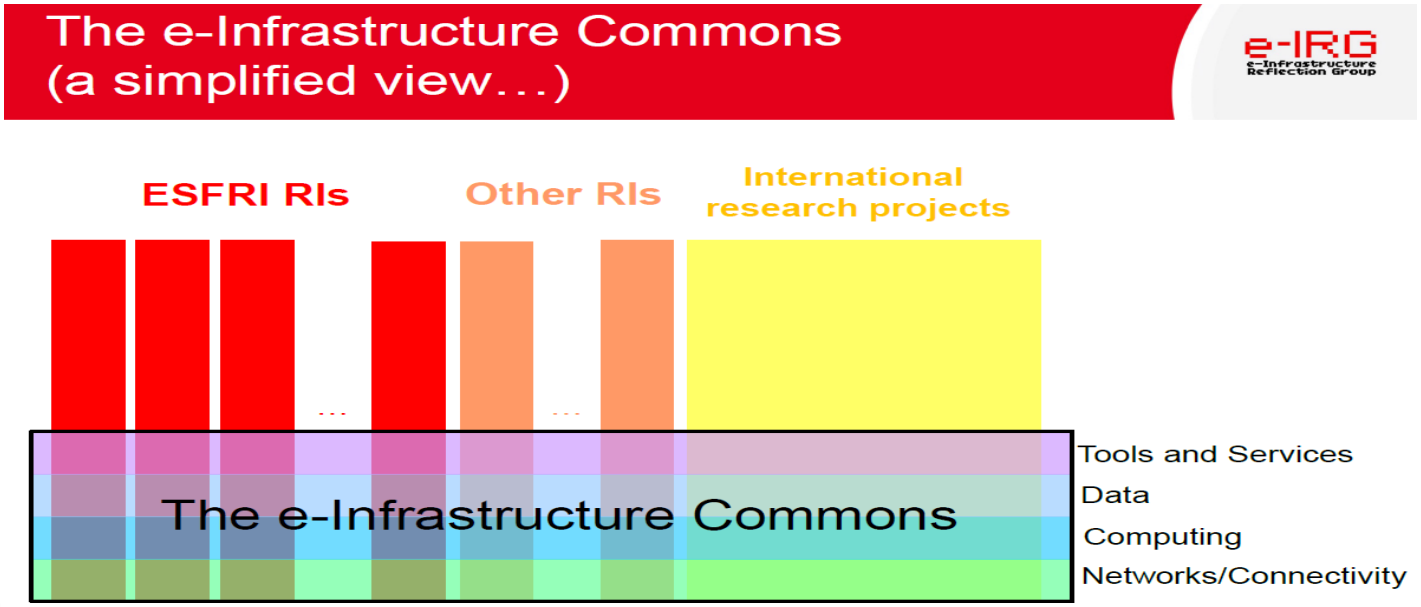




E-IRG The e-Infrastructure Commons

Establish a single e-Infrastructure Commons for knowledge, science and innovation …

- … provide a common interface towards users…
- … through a joint effort between users, strategic actors and e-infrastructure providers, to attain an ecosystem in which …
- … users enjoy the freedom to easily choose and use the services they need, so that …
- … they can focus on performing research, in (international) research collaborations.



RDA and global Data and Computing e-Infrastructure challenges: 20141211-12

Research Data and Computing infrastructures: Setting the Scene

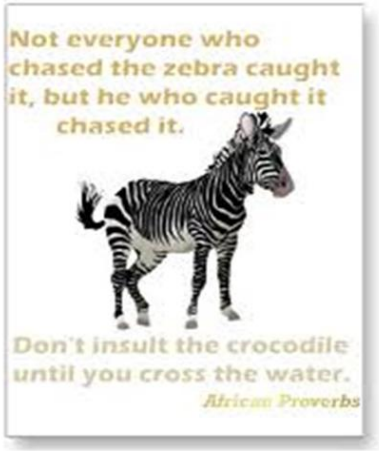
Panel … responding to key questions raised in the informal Ministerial meeting in Milan

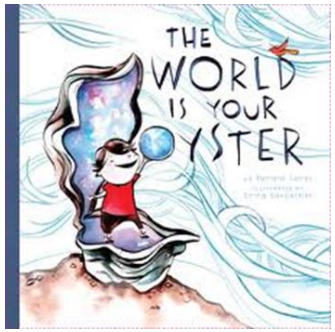
- *Research Infrastructures and cyber/ e-Infrastructures: what synergies?*
- *The data-driven computing challenges in the contexts of science*
- *Modernise science and education*

Promoting innovation in science: impact on growth & jobs

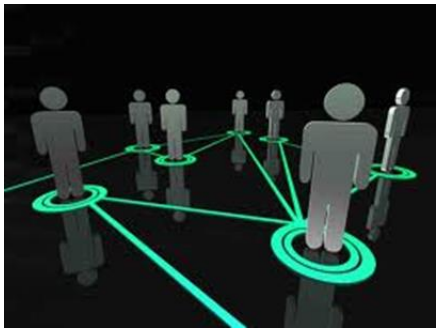
Opening science to society: education, training and communication

Enabling the integration of institutional, regional and national research capacity





Thank you



science & technology
Department:
Science and Technology
REPUBLIC OF SOUTH AFRICA

cjwright@csir.co.za

CSIR
our future through science