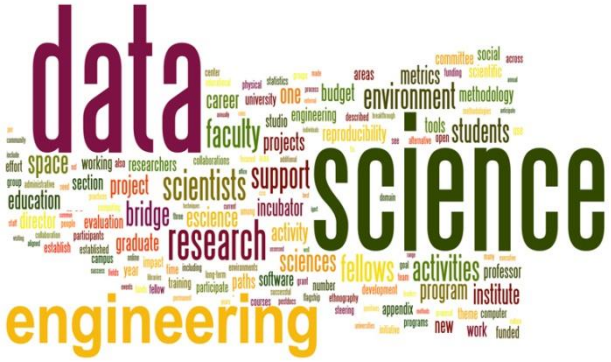
The 5th CSIR
CONFERENCE
IDEAS THAT WORK
8-9 October 2015 | CSIR ICC

Freedom to think: The Science of Data

Dr Quentin Williams

CSIR
our future through science

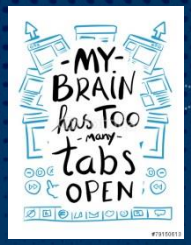
CELEBRATING 70 Years
Ideas that work

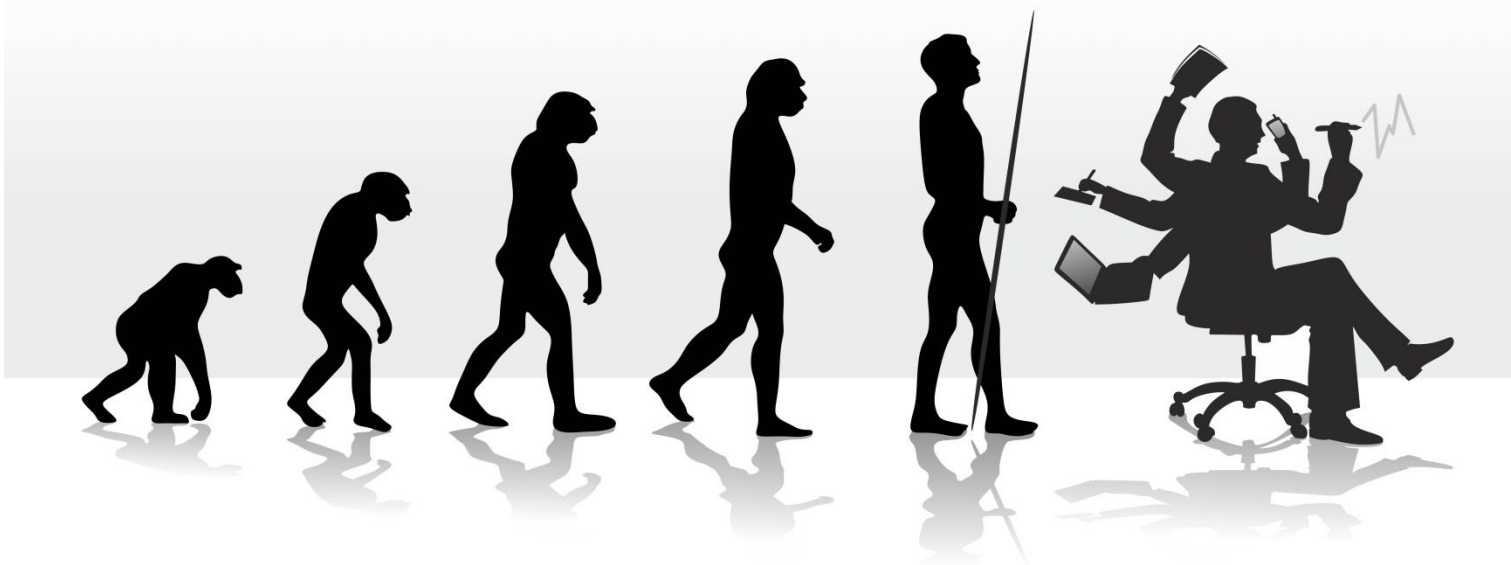


Content

- What is Data Science?
- The Importance of Data Science
- Data Enablement in South Africa
- The R&D behind the Science
- Success Stories
- Data Science in the CSIR

Freedom to think





Does all this influx of data allow us to operate more efficiently, effectively and competitively?

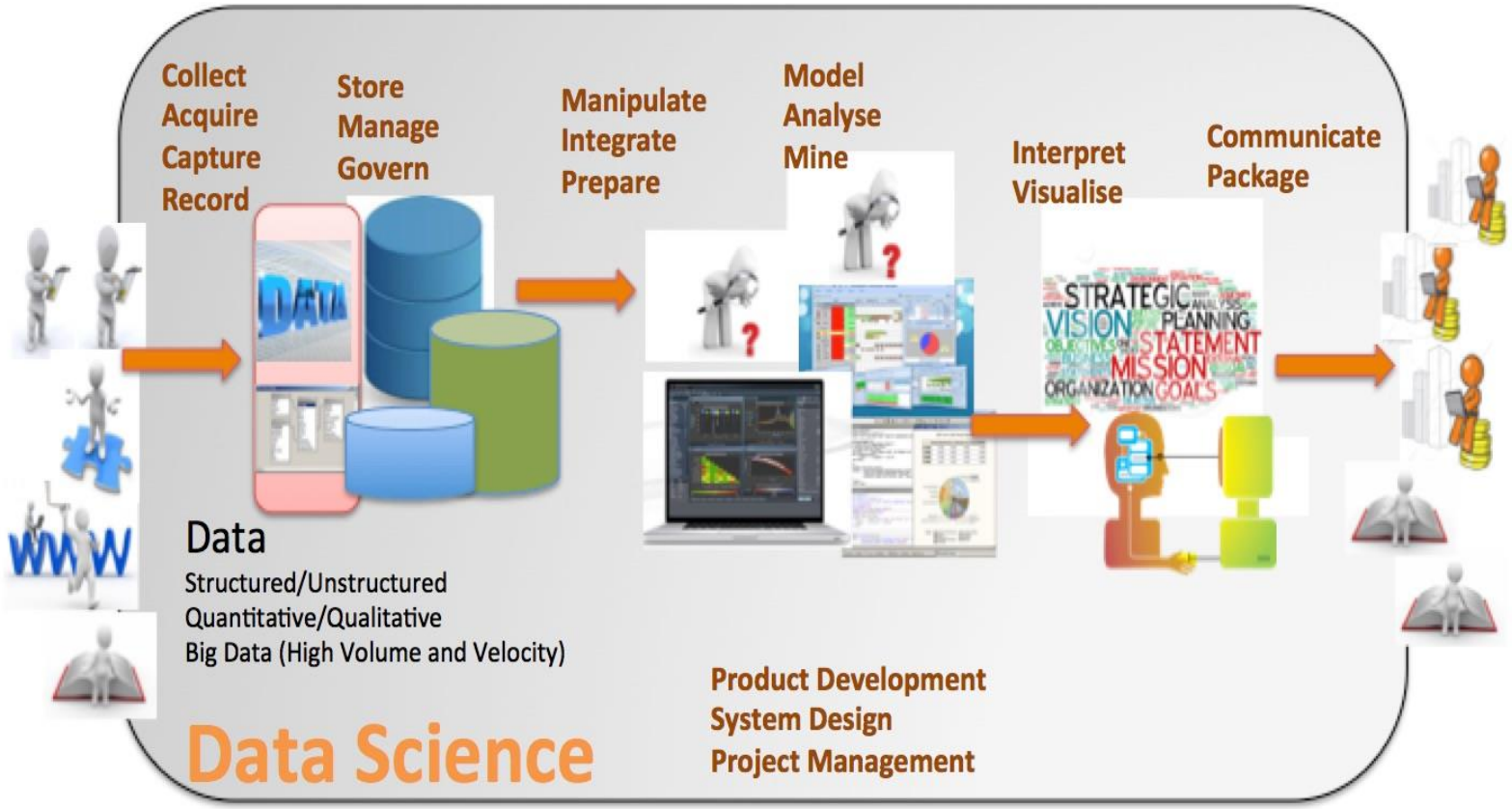
Are we more free or more confused?

What is Data Science?

Data science is a multi-disciplinary science with the main objective the extraction of value or development of products from available data.

Data science thus adopts a scope that deals end-to-end with everything data, including data acquisition, capturing and recording, integration, storage, manipulation, governance, analysis and mining, the visualisation and the final deployment to and use of the data products by the relevant stakeholders.

End-to-End Data



Decision-making

Before: Information gathered from hierarchical meetings: First group meetings, mid-management meetings, executive meetings

Now: Data collected from meetings and everything else, in real-time

Source: Robert Thomson, June 2014
"4 Strategic Planning Steps For Better Overall Decision Making", John Dillard, 2015

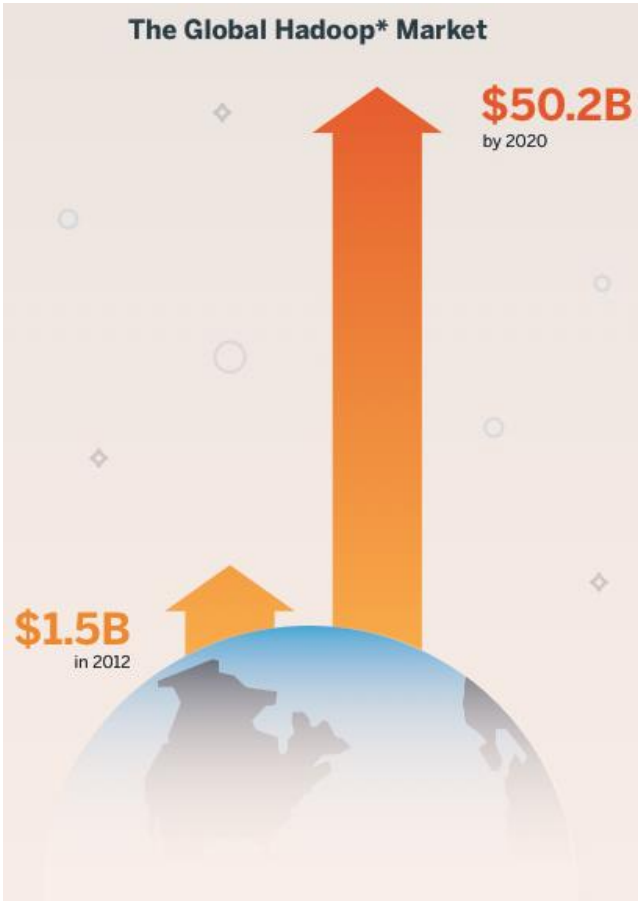
The Importance of Data Science

Data Science enables data discovery, helping you find new insights and ask questions you never knew to ask.

Integrate and analyse structured and unstructured data across all channels to better understand and discover, find new ways to compete, gain operational efficiencies and manage towards effectiveness.

Data Science allows the leverage and combination of data from multiple sources to gain a deeper understanding of interactions and interpret data to obtain faster insight and improve decision-making.

Data Enablement in South Africa



Accelerates the pace of discovery and new activities

- thus innovation and knowledge generation

Stimulates commerce and the economy

- thus new revenue streams

Addresses pressing NDP challenges in education, energy, health care, poverty, policy making

- thus better service delivery, efficient decision making, better resource management, skilled population

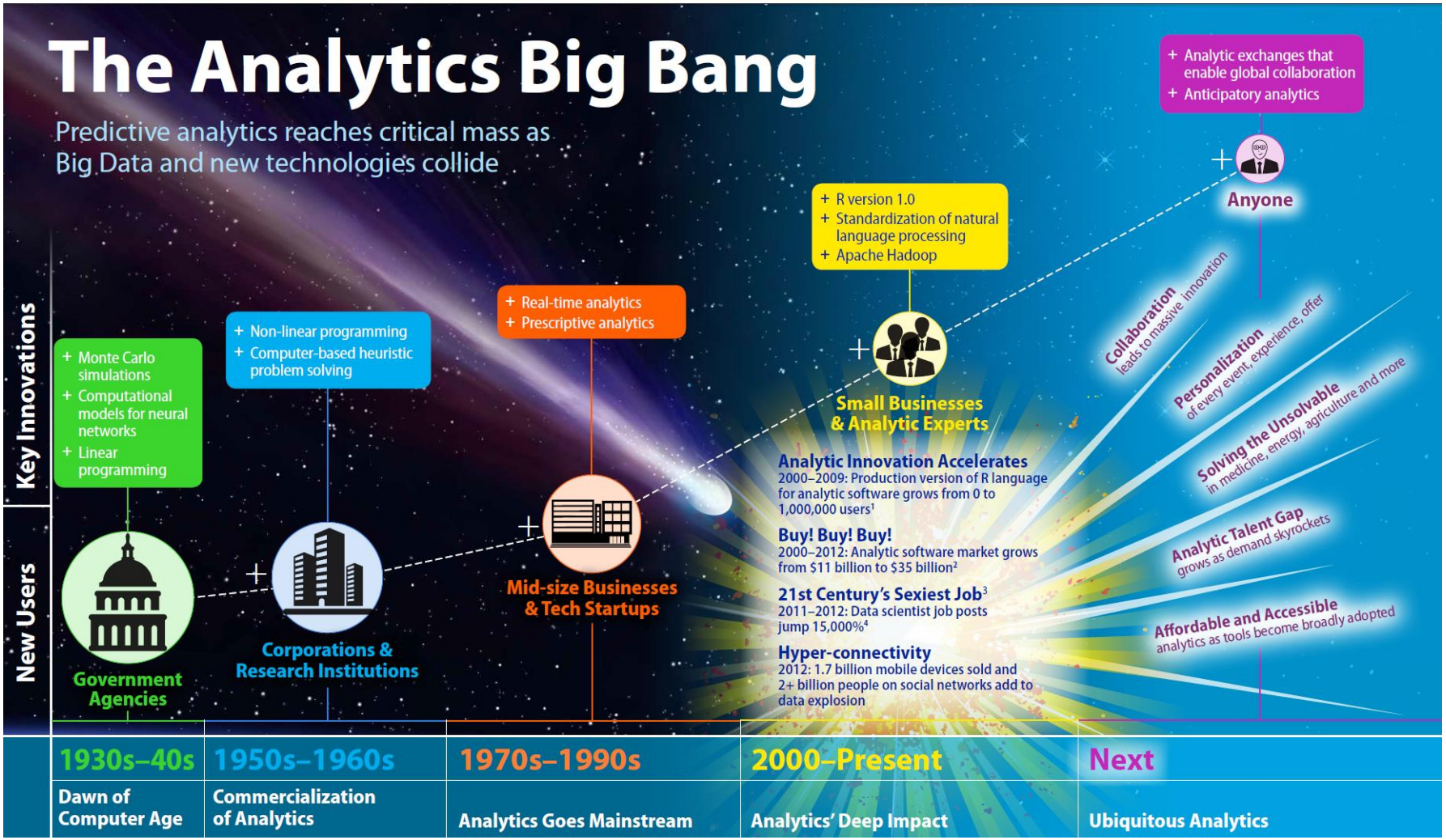
The 5th CSIR
CONFERENCE
IDEAS THAT WORK
8-9 October 2015 | CSIR ICC

The R&D behind the Science

CSIR
our future through science

CELEBRATING 70 Years
Ideas that work

Our Digital World

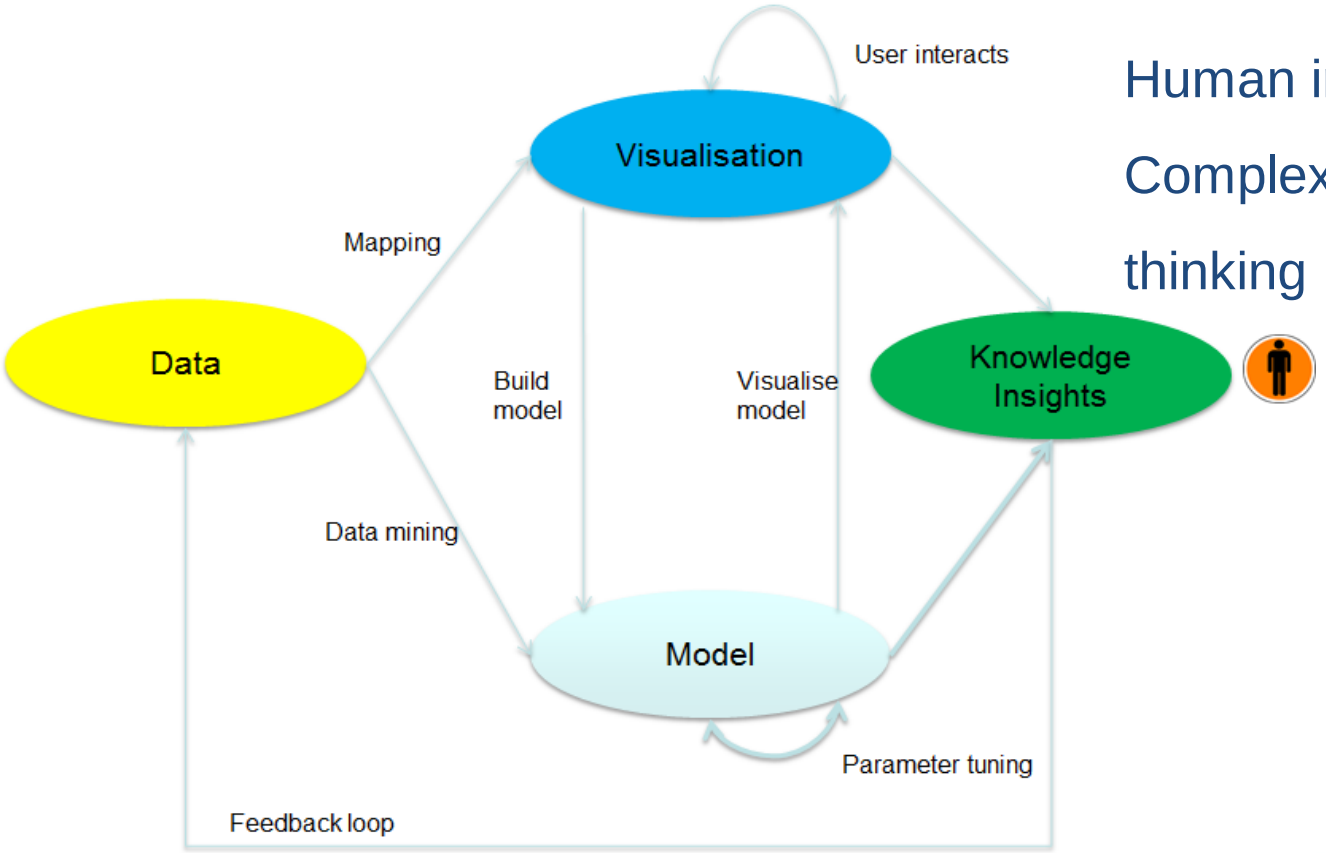


"The Analytics Big Bang", Fair Isaac Corporation, 2013

Data Science R&D

- Requires a multidisciplinary foundation that includes: Computing and software engineering, parallel computing, machine learning, statistics, optimisation, visualisation, and domain or product sense.
- Requires closing the gap between domain scientists and data scientists
- Requires the creation of open data repositories to allow data scientists to investigate the correlation of profiles using multi-source/sensor inputs.
- Requires applied and tested knowledge in data analysis pipeline:
  - Acquisition and recording, extraction, cleaning, and annotation
  - Analysis and modelling
  - Interpretation and insight

It really is about gaining Insight



Human in the loop: Complex and Creative thinking

Insight: understanding the causes of past events, predicting what might happen in the future

The 5th CSIR
CONFERENCE
IDEAS THAT WORK
8-9 October 2015 | CSIR ICC

Success Stories and CSIR examples

CSIR
our future through science

CELEBRATING 70 Years
Ideas that work

Global Success

- Zara is the flagship chain store of the Inditex group, the world's largest apparel retailer
- Co-founded by Amancio Ortega: from shop hand to 4th richest in the world (Forbes rich list 2015)
- Needs just one week to develop a new product and get it to stores, compared to the six-month industry average, and launches around 12 000 new designs each year
- Reduce lead times and react to new trends in a quicker way (“instant fashions”) using Data Science
- Store managers and employees are trained to talk to the customer and this data along with real-time sales numbers are instantly made available to the designers.
- The designers then update the product and send the final draft to the different manufacturers

The rise of Arsenal

Arsenal's 'secret' signing: club buys £2m revolutionary data company

- US-based StatDNA is expert in sports performance analytics
- Database helps guide Arsenal's scouting and game preparation

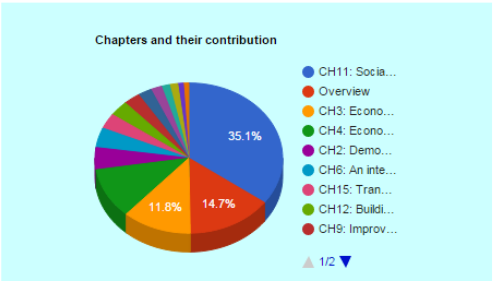
Source: The Guardian, 17 October, 2014

CSIR: SmartNDP Visualisation

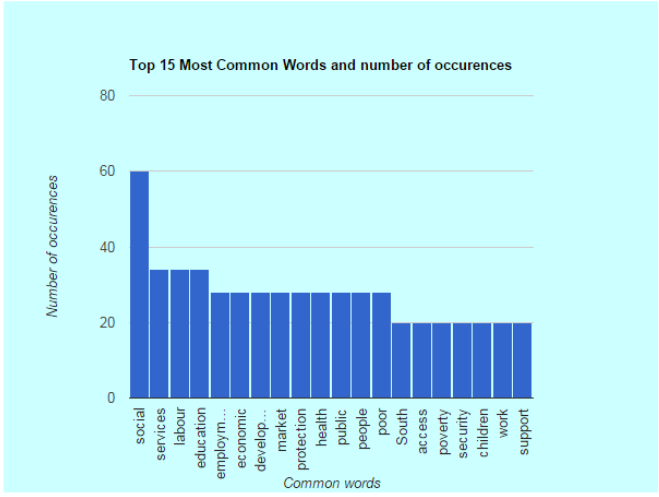
The 5th CSIR CONFERENCE
IDEAS THAT WORK
8-9 October 2015 | CSIR ICC

Digitised version of the National Development Plan (NDP) with the capability to track, monitor and evaluate projects and their contribution to the respective thematic objectives of the NDP

Chapter	Page	Text
Table of Contents	2	The foundation of social and economic development Introduction Access to basic electricity, water and sanitation, and public
Table of Contents	4	Introduction Evolution of social protection post-1994 Towards a broader social protection agenda Social security Developmental social welfare services Address the skills deficit in the social welfare sector Household food and nutrition security Creating a future social protection system Conclusion
Table of Contents	4	Introduction Vision 2030 Strengthen the criminal justice system: the seven-point plan Code of conduct National policing board Training for professionalism Demilitarising the police force An integrated approach to building safety A holistic view of safety and security Safety of women, the girl-child, children and the youth Rural safety Community participation in safety Conclusion
Overview	9	good-quality public services. All of these challenges are interlinked. Improved education, for example, will lead to higher employment and earnings, while more rapid economic growth will broaden opportunities for all and generate the resources required to improve education.
Overview	9	Mobilising society to support the plan, and exploring a social compact to reduce poverty and investment and employment. Conducting research on critical affecting long-term development. Advising government and social partners on implementing the plan. Working with relevant state agencies to report on the progress of the objectives.
Overview	10	Raising employment through faster economic growth Improving the quality of education, skills development and innovation Building the capability of the state to play a developmental, transformative role.
Overview	10	structural impediments, such as poor-quality education or spatial settlement patterns that exclude the majority. These are essential to achieving higher rates of investment and competi- tiveness, and expanding production and exports. Business, labour, communities and government will need to work together to achieve faster economic growth.
Overview	10	Political freedoms and human rights Social opportunities arising from education, health care, public transport and other public services Social security and safety nets An open society, transparency, disclosures and a culture of accountability Economic facilities, work, consumption, exchange, investment and production.
Overview	11	Direct and immediate measures to attack poverty The plan presents a long-term strategy to increase employment and broaden opportunities through education, vocational training and work experi- ence, public employment programmes, health and nutrition, public transport and access to informa- tion. While there are quick wins to be achieved in each of these areas, the strategies will take time to have a large-scale effect on poverty. To reduce the acute effects of poverty on millions of South Africans over the short term, the plan proposes to:
Overview	11	Introduce active labour market policies and incentives to grow employment, particularly for young people and in sectors employing relatively low-skilled people. Expand public employment programmes to 1 million participants by 2015 and 2 million by 2020. As the number of formal- and informal-sector jobs expands, public work programmes can be scaled down. Strengthen primary health-care services and broaden district-based health programmes, such as the community health worker and midwife programmes, and health education. Expand welfare services and public employ- ment schemes, enabling the state to service and support poor communities, particularly those with high levels of crime and violence. for Introduce a nutrition programme



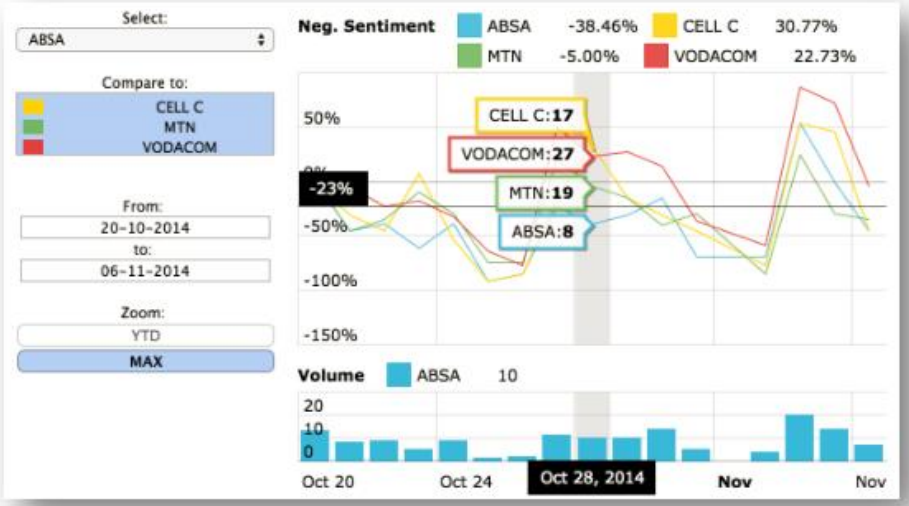
Page	Statistics
4	Introduction Vision 2030 Strengthen the criminal justice system : the seven-point plan Code of conduct National policing board Training for professionalism Demilitarising the police force An integrated approach to building safety A holistic view of safety and security Safety of women
11	Expand public employment programmes to 1 million participants by 2015 and 2 million by 2020
	Demographic trends What will South Africa look like in 2030 ? The country has reached a

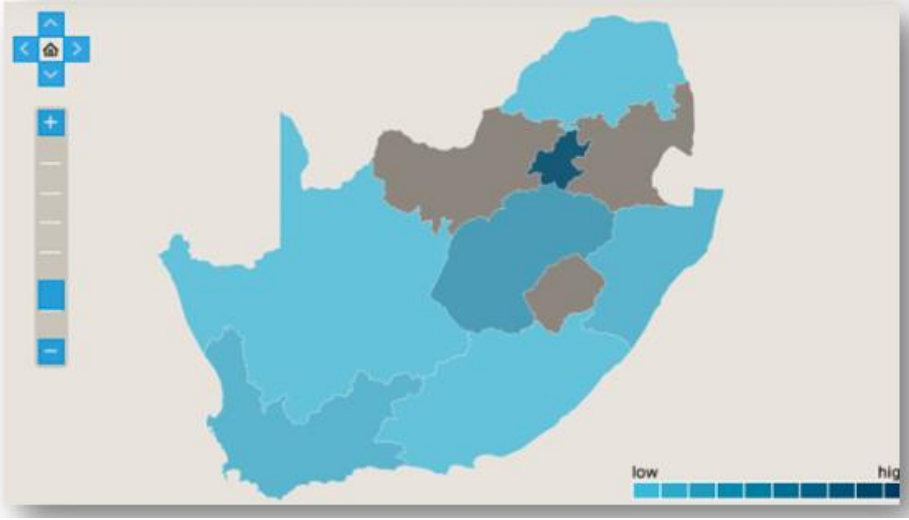


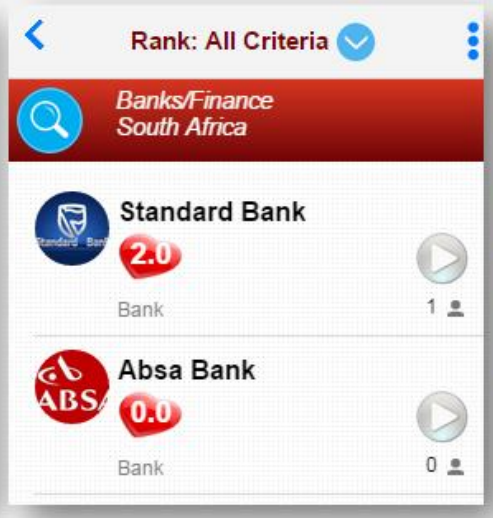
CSIR: Rateviz

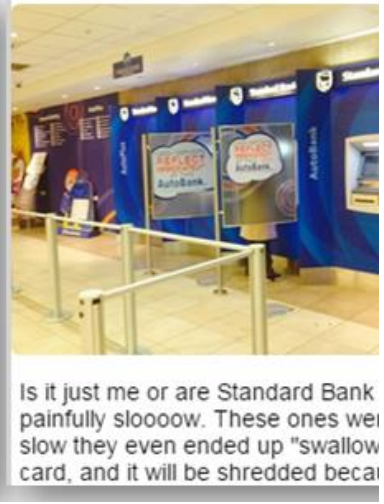
The 5th CSIR CONFERENCE
IDEAS THAT WORK
8-9 October 2015 | CSIR ICC

Rateviz: compare and analyze brands, visually



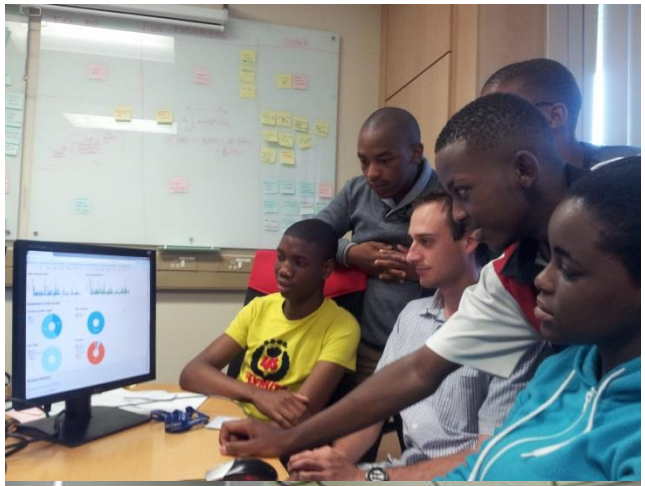


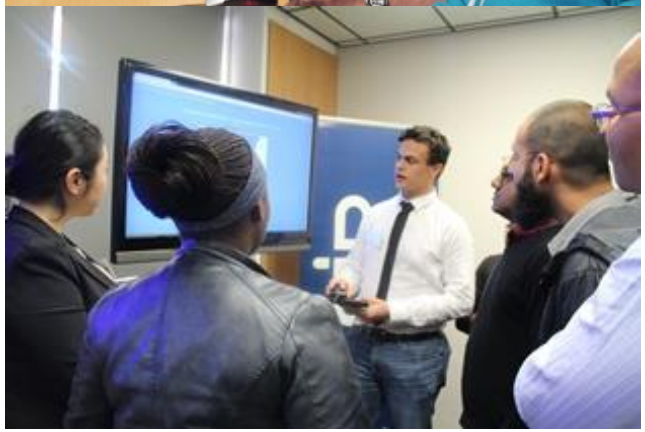




CSIR:
Data Science Skills Initiative

The 5th CSIR CONFERENCE
IDEAS THAT WORK
8-9 October 2015 | CSIR ICC

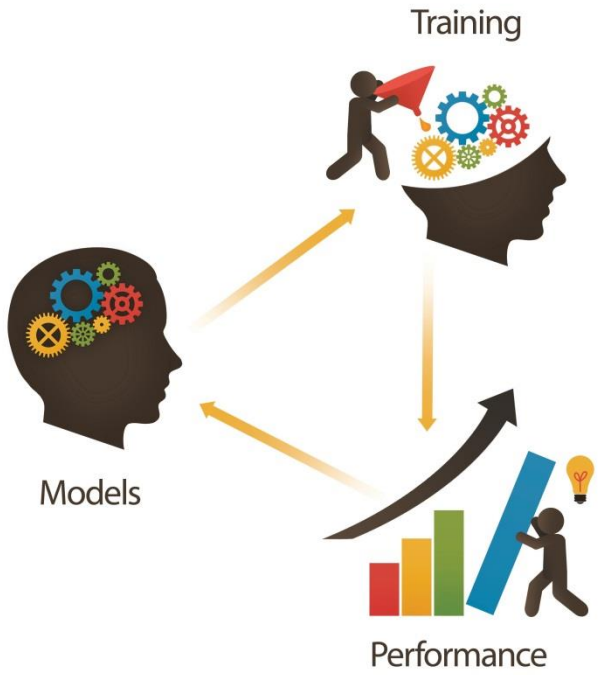




30 vacation trainees from 12 universities trained as part of the Data Science for Impact and Decision Enablement (DSIDE) programme to work on real-world problems presented by CSIR's stakeholders:

- Mentor-guided and learn-by-doing problem solving
- Take Business data: Analyse and Visualise
- Outputs from the students' projects are used by stakeholders for better decision-making

Conclusion: Freedom to think innovatively



New ways of operating, new ways of thinking, new ways of behaving:

- Individually: Manage your health, your appointments, thinking about others and the world, Improved insight
- Institutionally: More efficient (better operations), more effective (better utilisation of resources), more competitive (increase innovation), Improved decision-making

The 5th CSIR
CONFERENCE
IDEAS THAT WORK
8-9 October 2015 | CSIR ICC

Thank you

CSIR
our future through science

CELEBRATING 70 Years
Ideas that work