Institutional data curation implementation

1st African Digital Curation Conference

12 February 2008

Social science that makes a difference

HSRC
Human Sciences
Research Council

HSRC

The essence of the HSRC is evidence-based human and social science research that informs effective public policy debate to facilitate improvements in the living standards of South Africans

- Data producer
- Research areas
- Environment

- Policy analysis & Capacity Enhancement
- Knowledge Systems
- Social Aspects of HIV/AIDS Research Alliance
- Child, Youth, Family and Social Development
- Democracy and Governance
- Education, Science and Skills Development
- Social Aspects of HIV/AIDS and Health
- Service Delivery
- Education Quality Improvement
- Poverty, Employment and Growth

HSRC
Human Sciences
Research Council

Initial situation

- Lack of awareness
- No funding
- No emphasis on preservation
- No coherent strategy
- Limited expertise
- Limited technology
- Limited standards

Loss of data and documents

HSRC
Human Sciences
Research Council
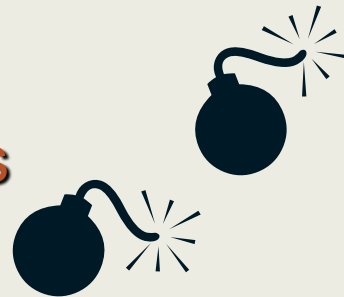
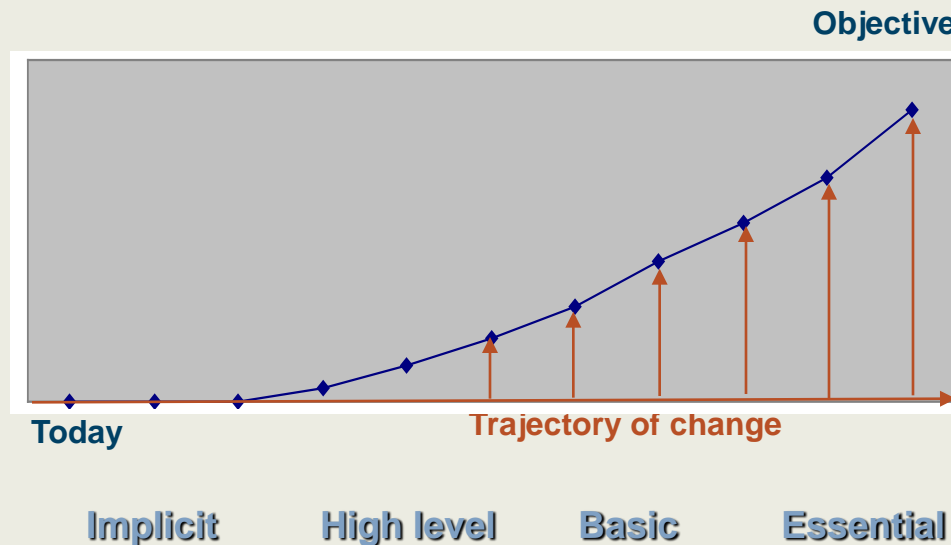# Critical success factors for implementation

Technology infrastructure

Organisational custody

Success

Processes

Resources: Funding Expertise

HSRC
Human Sciences
Research Council

A **successful implementation**

Organisational custody

Inter-organisational collaboration

- No awareness / Acknowledgement
- No commitment
- No policies

Objective

Today

Trajectory of change

Implicit　　**High level**　　**Basic**　　**Essential**

- Continuous, systematic management
- Comprehensive policy framework

A successful implementation

Technology infrastructure

Objective

•Heterogeneous

•Decentralised

•At risk

Today

Ad hoc

Reactive

Proactive
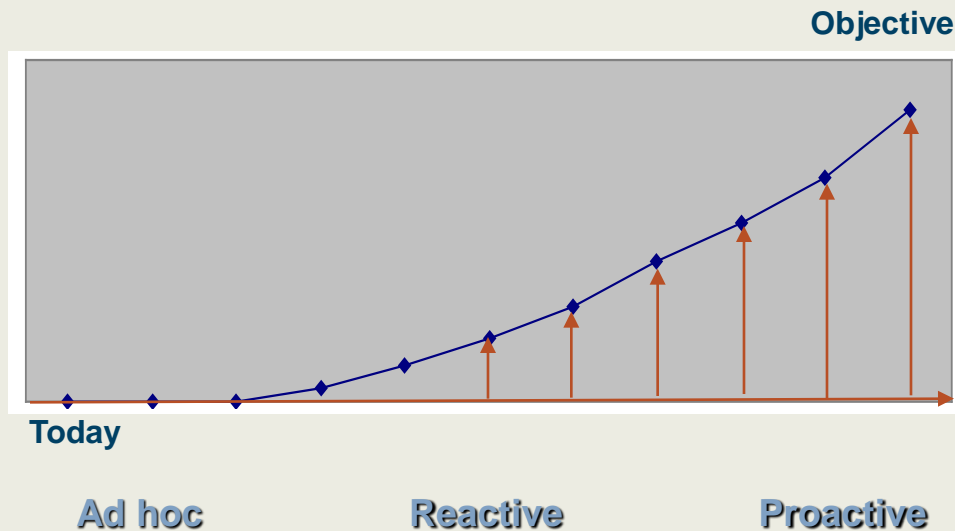
•Managed

•Integrated

•Secure

•Reliable

HSRC
Human Sciences
Research Council

A successful implementation

Resources - Funding, Expertise

Objective

- No funding
- Limited human resources
- Limited expertise
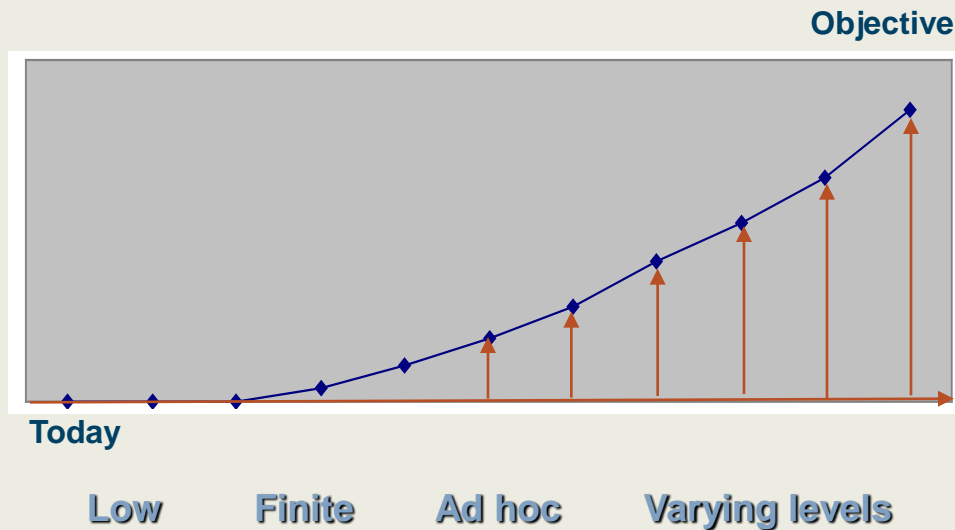
- Sustainable funding
- Well trained and sufficient human resources

Today

Low   Finite   Ad hoc   Varying levels

HSRC
Human Sciences
Research Council

A successful implementation

Processes

Objective

No / limited processes

Today

Ad hoc   **Disparate**   **Varying levels**

- **Well defined and integrated processes**
- **Executive ownership and support**

Social science that makes a difference

HSRC
Human Sciences
Research Council
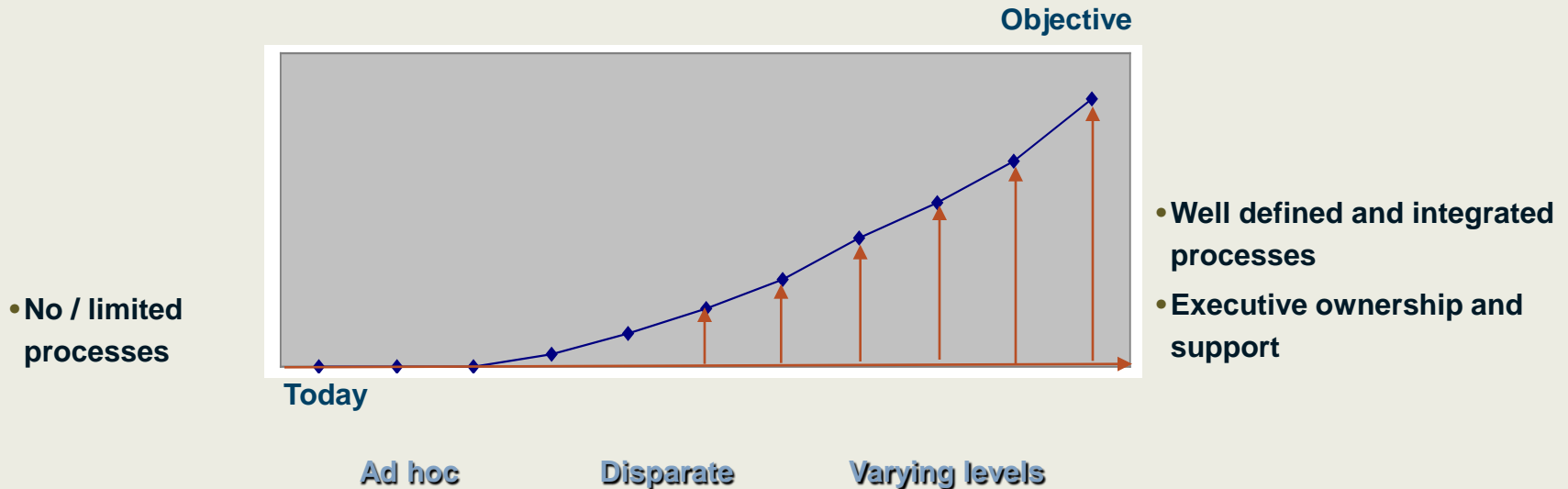
Organisational custody

- Ensure strategic alignment
- Obtain leadership (buy-in vs custody)
- Investigate the real effort and investment required and discuss with business executive
- Engage staff - management of change
  - Prove it works - pilot project
  - Road show - show benefit, consult
- Engage core group to assist with policy and standards development
- Create awareness for the need to disseminate of public data
- Create a culture of preserving and sharing data

Technology infrastructure

- Keep it simple
  - Create shared storage areas for researchers for final versions of data and documentation
  - Advise on standards
  - Create archiving and backup strategy - IT
  - Disseminate on organisational web portal
  - Download vs. software for online exploring
- Work with what you have
  - Make optimal use of existing systems
  - Make use of open source software
  - Choose proven technologies
  - Stick to essentials
- Partner with other organisations

Social science that makes a difference

HSRC
Human Sciences
Research Council

Resources - Funding, Expertise

- Funding
    - Funding of pilot project
    - Obtain sustainable funding
- Expertise
    - Make use of existing expertise
    - Encourage appointment of high-level data managers
    - Develop expertise by exposure to research and trends (international) and collaboration
    - Learn from others who have already done successful implementations

Social science that makes a difference

HSRC
Human Sciences
Research Council

Processes

- Engage on standards

- Streamline process flow from data management to curation

- Distribute responsibilities

- Promote a culture of sharing

Implementation plan

Lifecycle of research data

External: Web Portal Dissemination software

IT

Project registration → **Internal: RMS**

Data creation
Survey/measurement
Data entry
Data checking and cleaning

Secondary data

Dissemination

IT

Data Management, Library

Raw data

Data analysis
Analysis
Derived data creation: imputing missing data, new variables, weights
Linkage to other data
Creation of documentation

Preservation: Ingest, Storage

Prepare dissemination formats

Archive data with additional metadata and documentation (preservation format)

Additional derived/linked data

Data Management, Library, IT

Curation appraisal

Appraisal Committee

EDs, Data Management

Save master/final version to central location with minimum metadata

Researcher

Novell file structure / DMS

Data management & curation standards

Social science that makes a difference

HSRC
Human Sciences
Research Council

Success is not guaranteed by technology.

Success

It is about what the business and people do with the technology that creates benefit, efficiency and competitive advantage.

HSRC
Human Sciences
Research Council