Data Science: A New Discipline or an Agglomeration of Skills?

- Professor Gobinda Chowdhury
- Professor, iSchool@northumbria, Head, Dept. of Mathematics & Information

- 24/11/2014

Data Science

- Data Science is a technical field that represents "*an evolution from the business or data analyst role*" (IBM)
- "*the field of data science is emerging at the intersection of the fields of social science and statistics, information and computer science, and design*." (UC Berkeley)

northumbria
UNIVERSITY NEWCASTLE

Data Science today is where Computer Science was 50 years ago

- In one sense today the emerging field of Data Science is facing the same identity challenges that the field of computer science faced for nearly 20 years in the middle of the last century:
- *"The discipline of computer science was born in the early 1940s with the confluence of algorithm theory, mathematical logic, and the invention of the stored-program electronic computer.*

northumbria UNIVERSITY NEWCASTLE

Data Science Skills

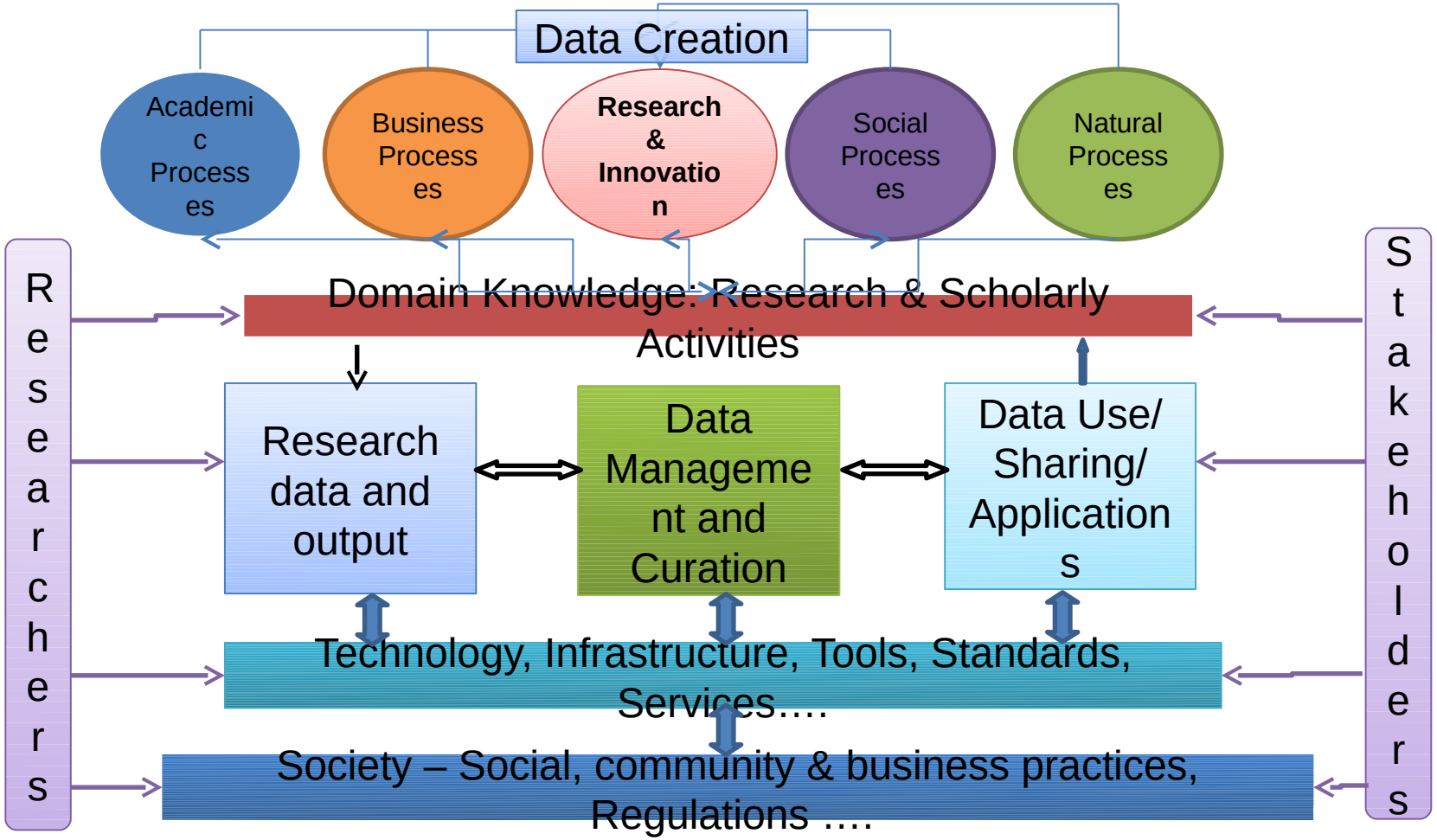
- a set of soft skills (e.g., user studies and social study skills to understand the scope and dimensions of the information behind the data collected)
- a set of technical and computational skills (e.g. e-infrastructure, software engineering and systems studies)
- Research application domains (e.g., health, transportation, environment,…) and

northumbria
UNIVERSITY NEWCASTLE

Managing Data: Research and Innovation

- Managing Information Vs. Managing data
  - Challenges
- Different Perspectives
  - Researchers
  - Technology and infrastructure service providers
  - Information service professionals

northumbria UNIVERSITY NEWCASTLE

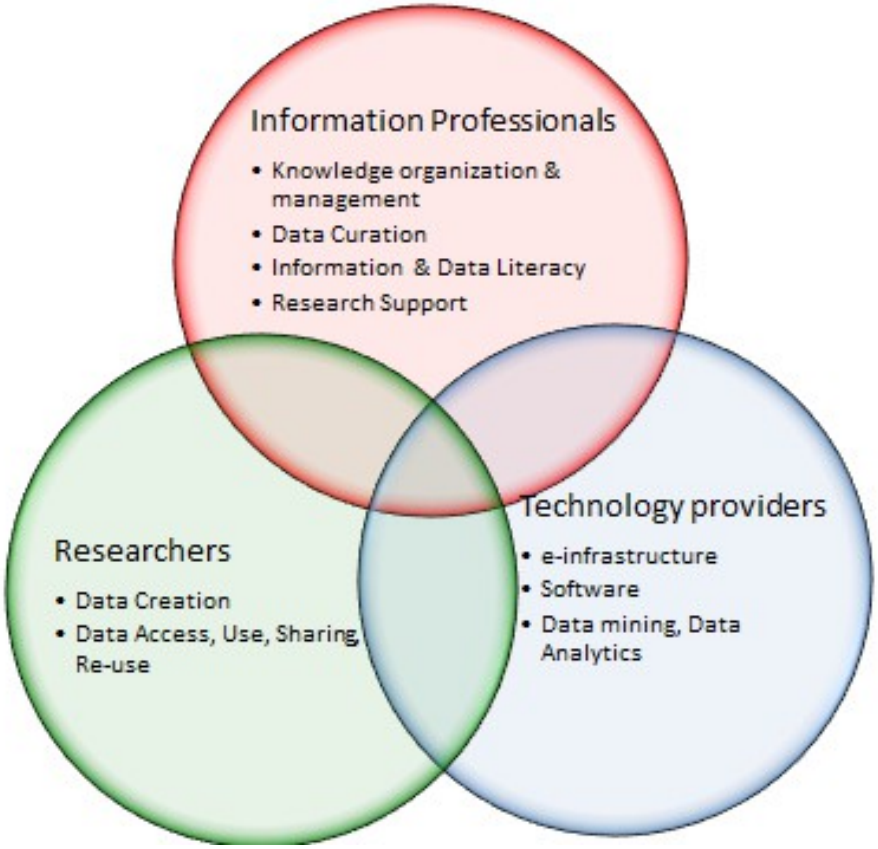

Data Creation
Academic Processes
Business Processes
Research & Innovation
Social Processes
Natural Processes
Researchers
Domain Knowledge: Research & Scholarly Activities
Research data and output
Data Management and Curation
Data Use/ Sharing/ Applications
Stakeholders
Technology, Infrastructure, Tools, Standards, Services….
Society – Social, community & business practices, Regulations ….

Research Data Management

- Why and how is it different from IM?
- IM:
  - Building block is words
  - Word meaning 9semantics0 is established
  - Word relations and hence meaning is controlled by syntax and context
- DM:
  - Building block is Data – mainly alphanumeric characters

northumbria UNIVERSITY NEWCASTLE

So, what skills do we need?



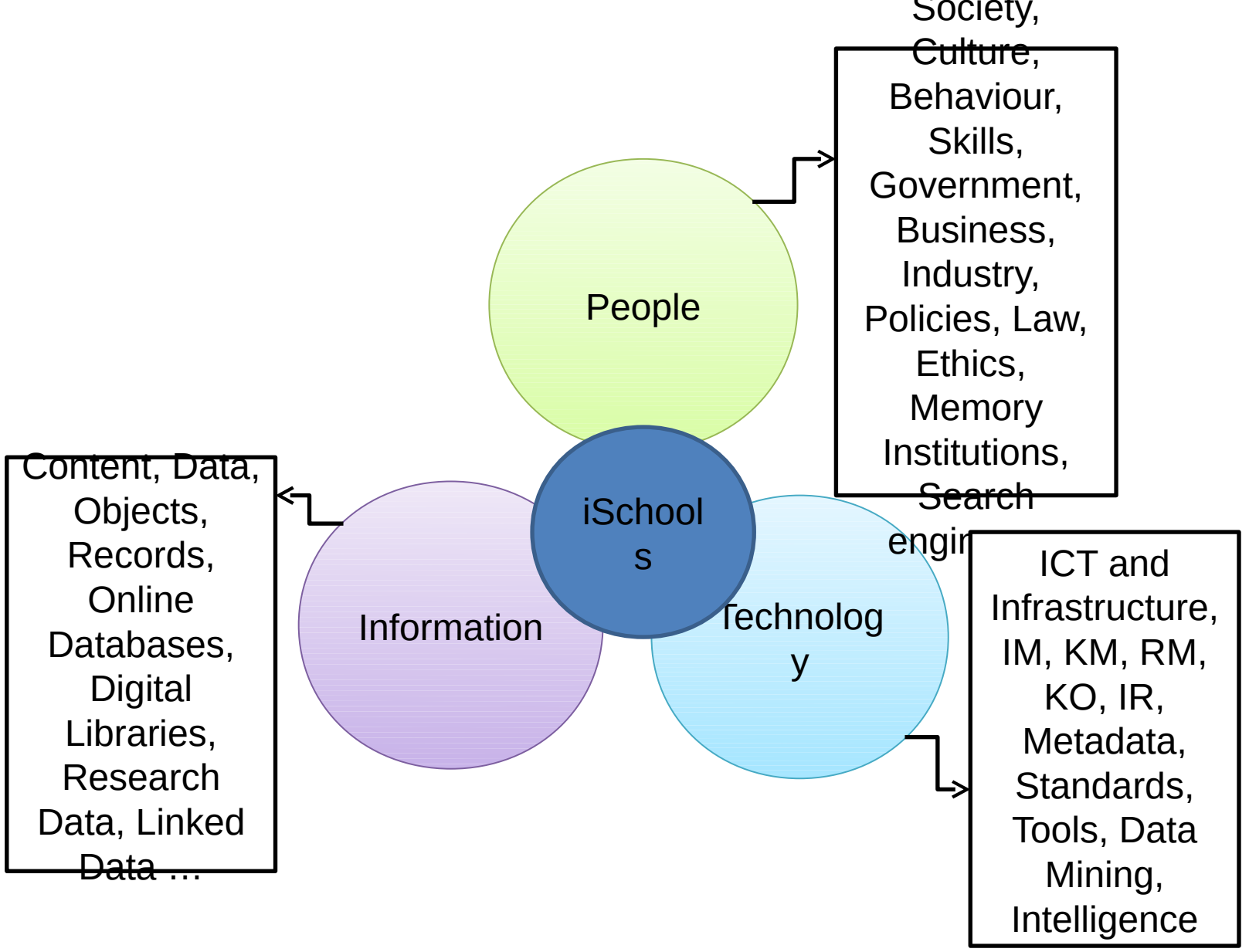
Data management: what skills do we need

- About data:
  - nature of data; meaning of data; how data is produced; how data is stored; how data is linked; how data is used – application rules …
- About technology
  - What infrastructure, technologies, tools, standards?
- About people and disciplines
  - Who produces and who uses data? why and how? how the user and community behaviour

Northumbria UNIVERSITY NEWCASTLE

What Can the iSchools do?

- What skills set do the iSchools offer?
- How can they be integrated and re-packaged?


People
Information
Technology
iSchools
Society, Culture, Behaviour, Skills, Government, Business, Industry, Policies, Law, Ethics, Memory Institutions, Search engin
Content, Data, Objects, Records, Online Databases, Digital Libraries, Research Data, Linked Data ...
ICT and Infrastructure, IM, KM, RM, KO, IR, Metadata, Standards, Tools, Data Mining, Intelligence

Document4 - Microsoft Word

File Home Insert Page Layout References Mailings Review View Acrobat

Paste Cut Copy Format Painter

Calibri (Body) 11

Clipboard Font Paragraph Styles Editing

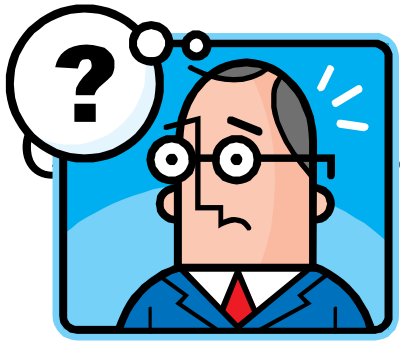
AaBbCcDc ¶ Normal AaBbCcDc ¶ No Spaci... AaBbC Heading 1 AaBbCc Heading 2 AaB Title AaBbCc. Subtitle

Change Styles

Find Replace Select

Page: 1 of 1 Words: 82 100%

northumbria UNIVERSITY NEWCASTLE

…. And Now

- Thanks for listening
-