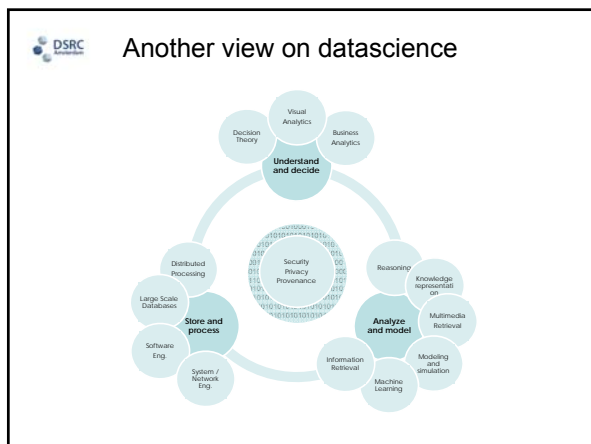
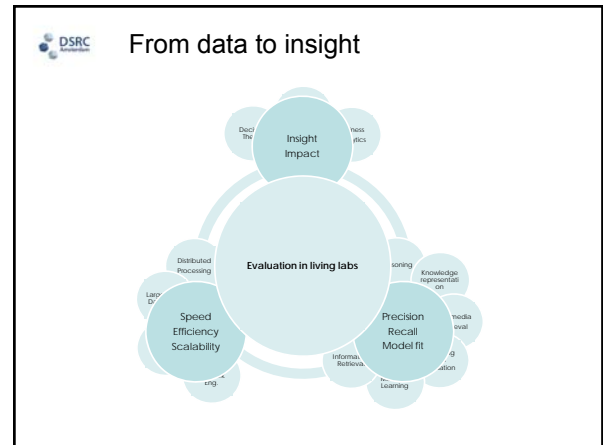
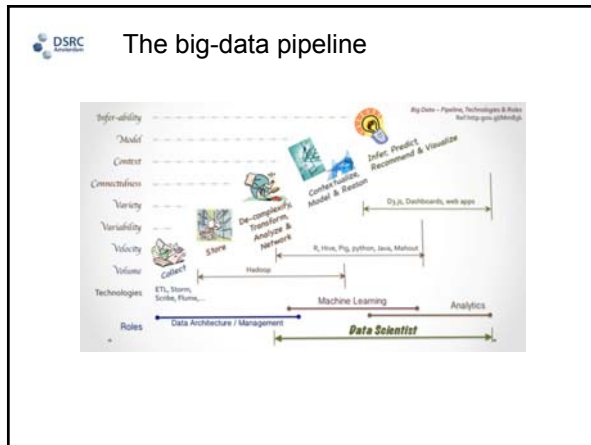
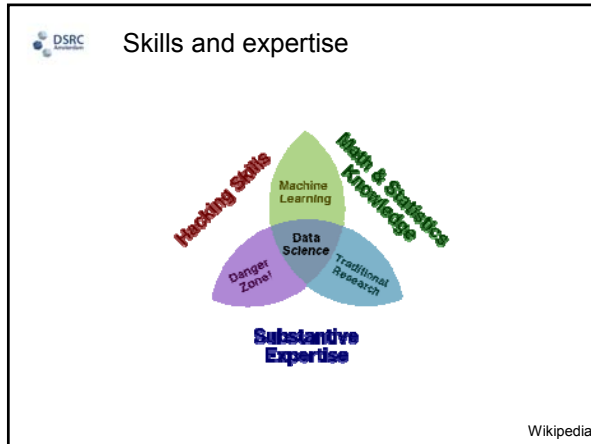


- Why a data science approach?
- Hal Varian (**Google's chief economist**)
 - "The ability to take data — to be able to understand it, to process it, to extract value from it, to visualize it, to communicate it — that's going to be a hugely important skill in the next decades".



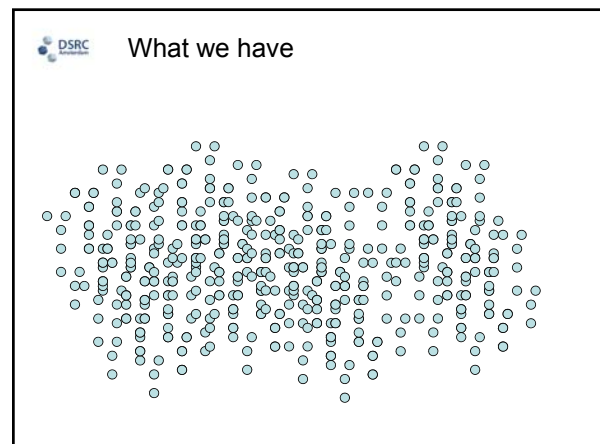
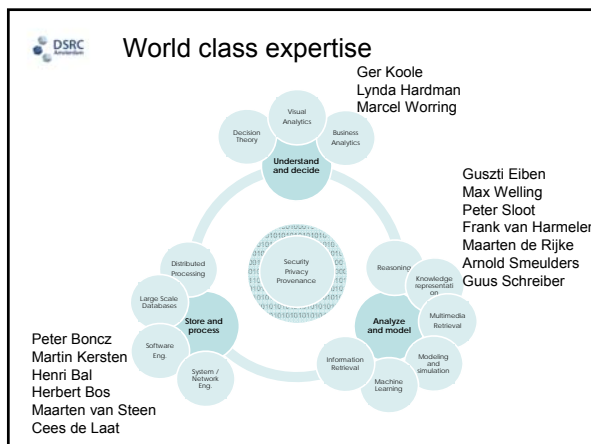
- DSRC Data Science
- Characteristics
 - All are connected
 - All driven by data and its use
 - A holistic approach is needed
 - Our answer
 - The Data Science Research Center



- ### Ambitions and approach
- Ambitions**
 - First class academic and applied research
 - Focus on new ICT development
 - Education
 - BSc
 - MSc
 - PDEng
 - Entrepreneurship
 - Societal impact
 - Approach**
 - Data-intensive internal/external collaborations
 - Education
 - Start: AI data science track
 - Living labs
 - ACE Venture Labs
 - International collaborations

- ### Basic Facts of the DSRC
- Founded Fall 2013
 - Four academic partners
 - Built around multiple proven research strengths in which we are world leaders
-

- ### “Applications: Big City Data” broadly conceived
- Creative industry**
 - Cultural applications
 - Broadcasting
 - Health care**
 - Customer optimization processes
 - Supporting elderly in their homes
 - Finance**
 - Learning and prediction of customer behavior
 - Scientific Data**
 - Simulating epidemics
 - Biodiversity analysis
 - Smart city / Smart citizen**
 - Events and news



DSRC What we want?

Insight

Visualization is the answer?

DSRC Some more examples

DSRC What is insight? North CG&A, 2006

Complex
Insight is complex, involving all or large amounts of the given data in a synergistic way, not simply individual data values.

Deep
Insight builds up over time, accumulating and building on itself to create depth often generating further questions and, hence, further insight.

Relevant
Insight is deeply embedded in the data domain, connecting the data to existing domain knowledge and giving it relevant meaning going beyond dry data analysis, to relevant domain impact.

Qualitative
Insight is not exact, can be uncertain and subjective, and can have multiple levels of resolution.

Unexpected
Insight is often unpredictable, serendipitous, and creative.

DSRC Examples

DSRC Some example visualizations (from show me the data)

DSRC Finding patterns in complex data

• Visual Analytics
– Combine the power of computer and human

- Compute power
- Storage capacity

- Flexibility
- Creativity
- Expert knowledge

DSRC Visual Analytics

Visual Analytics is the science of analytical reasoning facilitated by interactive visual interfaces

"Detect the expected, discover the unexpected"

DSRC The basis: Concept detection

Humans can do this very well, but not for 100.000 images

Visual examples
Positive Negative

Unknown images → Learn model → Score of presence → ranking

DSRC Analytics

- What is the best known Analytic tool?

Yes the Spreadsheet

DSRC Multimedia Analytics Process

DSRC Analytics

Fischer et al. TVCG 2010.

So how about multimedia?

DSRC Conceptualizing the problem

- Define the initial categories of interest

DSRC Framing the problem

Filtering to obtain reasonable size

Refining specific concepts

DSRC Towards picture insight Worring and Koelma

Pivot Tables for Multimedia data

Sort and weight

Decompose

Aggregate Visualization

Flexible multimodal summaries

DSRC Analytics: Pivot Tables

Filter variables: Any faceted combination

Column variables: Sort and Weight

Category	Type	Integer	Metadata
VALUE	VALUE	VALUE	VALUE

Row variables: Decompose

Column aggregation

DSRC Conclusion: Insight

Complex

Deep

Relevant

Incremental categorization as filters and variables

How to make it more intelligent?

categories with user annotations

patterns / outliers

Incremental categorization

Multimodal categorization

Multimedia pivot table

User in context

Work with domain

DSRC Possible insights

- A new category
 - Sub-category of an existing category
 - Grouping / generalization of categories
 - Other
- Data characteristics
 - Individual elements
 - Between individual elements
 - Individual category
 - Between categories
 - Of specific metadata values

DSRC The things we need

- Visualizations
- Loop between human and machine
- A lot of automatic analysis
 - Information Retrieval
 - Computer Vision
 - Machine Learning
 - Knowledge representations
 -
- High performance computing
-

