ACM Learning Center
http://learning.acm.org

- 1,300+ trusted technical books and videos by leading publishers including O'Reilly

- ACM Tech Packs on big current computing topics: Annotated Bibliographies compiled by subject experts

- Online courses with assessments and mentoring (coming soon!

- ACM Learning Paths providing unique, accessible entry points into popular languages such as Python and Ruby

Association for
Computing Machinery

Advancing Computing as a Science & Profession

Barry Devlin

Founder and Principal
9sight Consulting, www.9sight.com

Dr. Barry Devlin is a founder of the data warehousing industry and among the foremost authorities worldwide on business intelligence (BI) and beyond. He is a widely respected consultant, lecturer and author of "Data Warehouse—from Architecture to Implementation". Barry has 30 years of experience in the IT industry, previously with IBM, as an architect, consultant, manager and software evangelist.

As founder and principal of 9sight Consulting (www.9sight.com), Barry provides strategic consulting and thought-leadership to buyers and vendors of BI solutions. He is currently developing a new architectural model for fully consistent business support—from informational to operational and collaborative—Business Integrated Insight (BI^2). Based in Cape Town, South Africa, Barry's knowledge and expertise are in demand both locally and internationally.

Email: barry@9sight.com (preferred contact method)
Phone: (S. Africa cell)   +27 71 557 7479

Ankur Teredesai

Associate Professor of Computer Science University of Washington, Tacoma; SIGKDD

Ankur M. Teredesai's research interests include Data Science for the Web. Ankur has worked on a variety of Big Data problems for internet monetization and advertising, trust-enhanced social recommendation systems, handwritten zip-code recognition, novelty detection in video data streams, and stream data management. His work has appeared in numerous publications, as well as international conferences and research panels.

Ankur is currently the Information Officer for ACM SIGKDD, a worldwide organization of data science and big data professionals from both academia and industry. Ankur has been a technical advisor for several large data warehouse and data analytics groups at Microsoft, Apollo Data Technologies (now MethodCare), and Davai.com. He currently serves as a data science advisor for a prominent advertising big data startup nPario.com.

2012 - Big Data: End of the World or End of BI?

ACM Learning Webinar

June 2012

Dr Barry Devlin

Founder & Principal
9sight Consulting
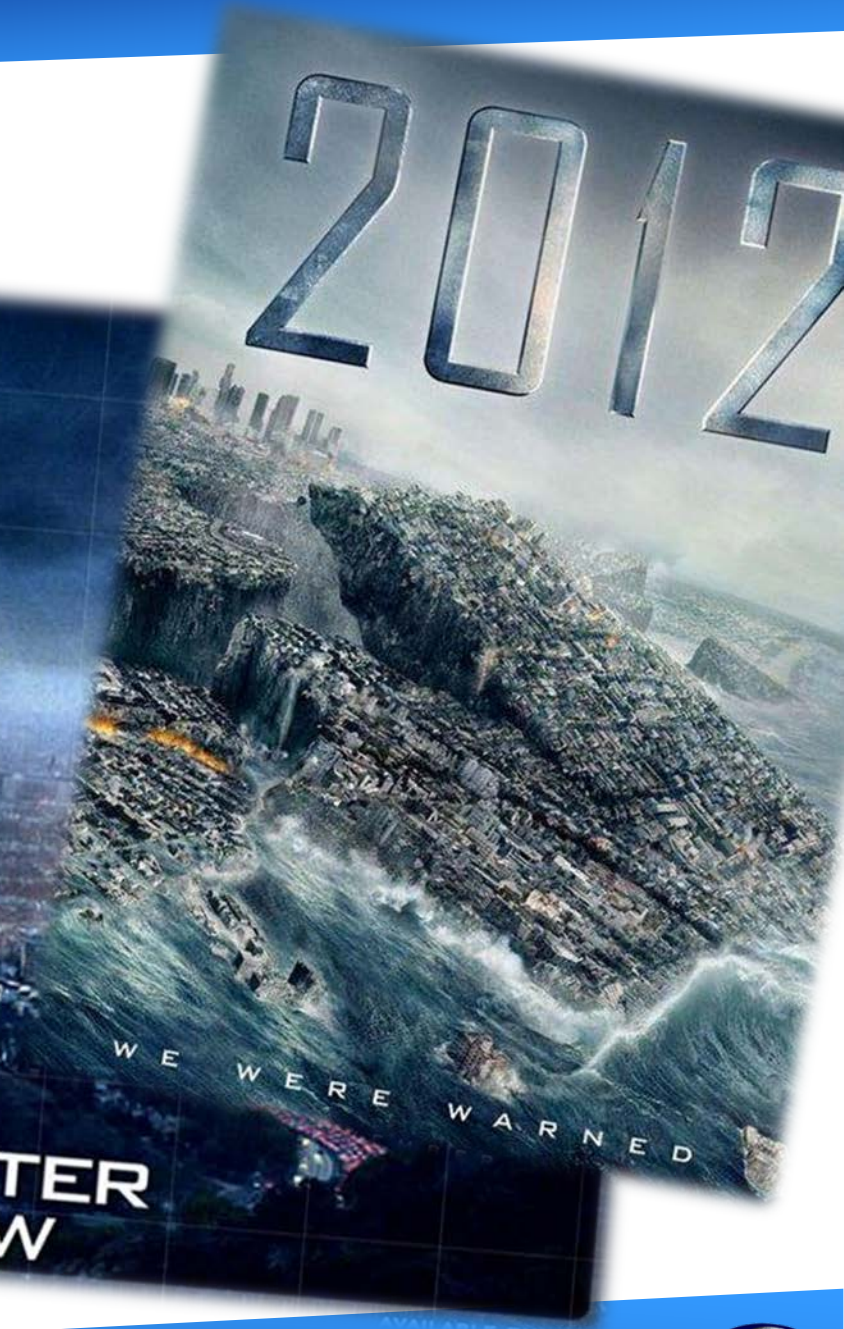
Association for
Computing Machinery

Advancing Computing as a Science & Profession

A brief history of...

The Original Data Warehouse Architecture (1988)

- Based on internal work in IBM Europe from 1985 on

- *"Business Data Warehouse (BDW)… is the single logical storehouse of all the information used to report on the business… In relational terms, the end user is presented with a view / number of views that contain the accessed data… The user thus sees a set of tables containing only needed columns, although these may have been obtained from … different tables"*

- Raw & enhanced, detailed & summary, public & personal data all within a single component

- *"An architecture for a business and information system"*, B. A. Devlin, P. T. Murphy, IBM Systems Journal, Vol .27, No. 1, (1988)
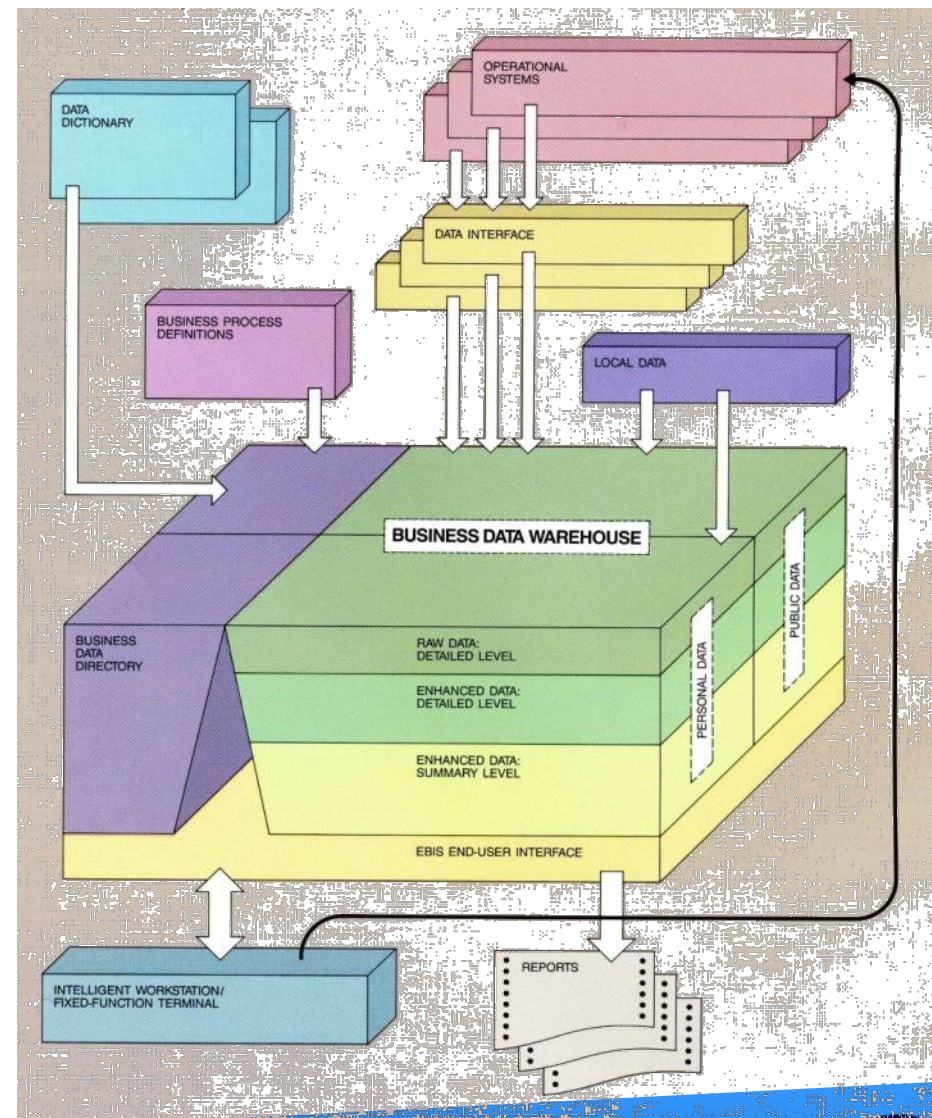
The layered Data Warehouse since the early '90s

Operational / Informational split

Two layers within the DW

1. Enterprise data warehouse
    – Reconciled data

2. Data marts
    – What the users need

- Characteristics
    – Vertical and horizontal segmentation of information
    – Separate metadata
    – Hard data only
    – Unidirectional data flow

- Well architected!

Data marts (Informational Apps)

Metadata

Enterprise data warehouse

Operational systems
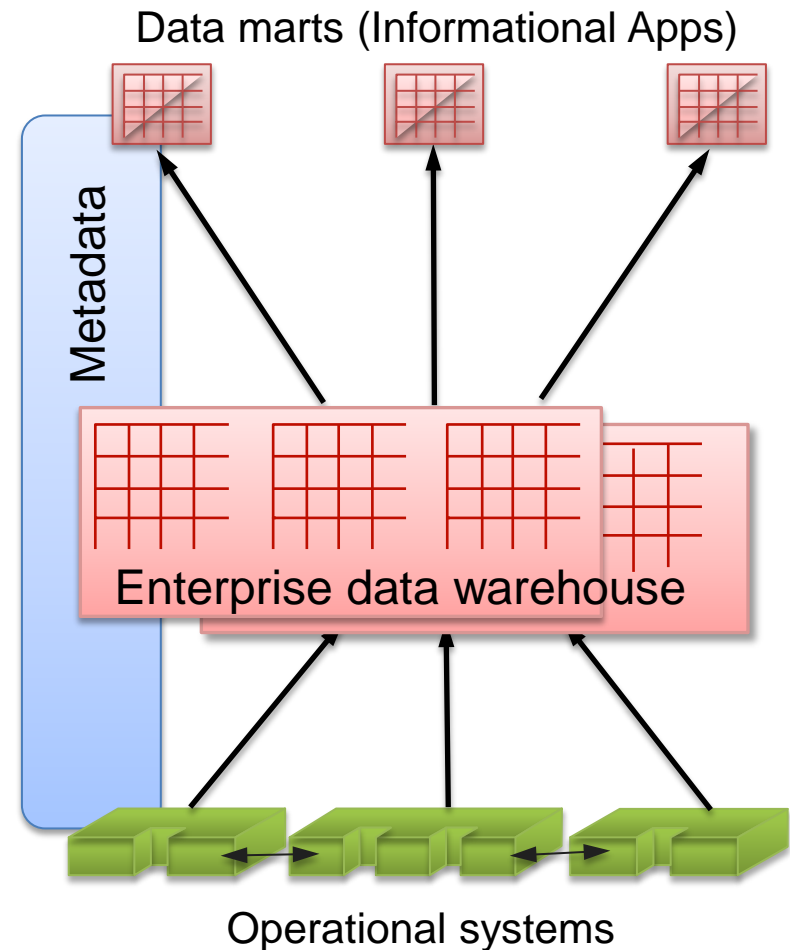
The four ancient postulates of data warehousing.

Postulate 1 (1970s):

Operational and informational environments should be separated for both business and technical reasons.

Postulate 2 (1980s):

A data warehouse is the only way to obtain a dependable, integrated view of the business.

Postulate 3 (1980s):

The data warehouse is the only possible instantiation of the full enterprise data model.

Postulate 4 (1990s):

*A layered data warehouse is necessary for speedy and reliable query performance .*

See: Devlin, B. *"Business Integrated Insight (BI^2): Reinventing enterprise information management"*, (2009), www.9sight.com/resources.htm

Explosion of DW components – mid-'90s onwards

- Changing business needs lead to addition of new components:

  - Operational data store
    - Near real-time
  - Data Staging Area
    - Data conditioning
  - More types & numbers of marts
    - Marts fed from marts
    - Spreadsheets a real issue
  - Independent data marts
    - Bypassing EDW, often very large
  - Bidirectional data flows
    - Merging op. / info. needs
  - Federated / Virtual access
    - Accessing operational systems

- Hmmm… not so well architected!

Data marts, cubes, spreadsheets, independent data marts, etc.

Metadata

Mashups, Federated query

Enterprise data warehouse

Data Staging Area

Operational data store

Operational systems and more
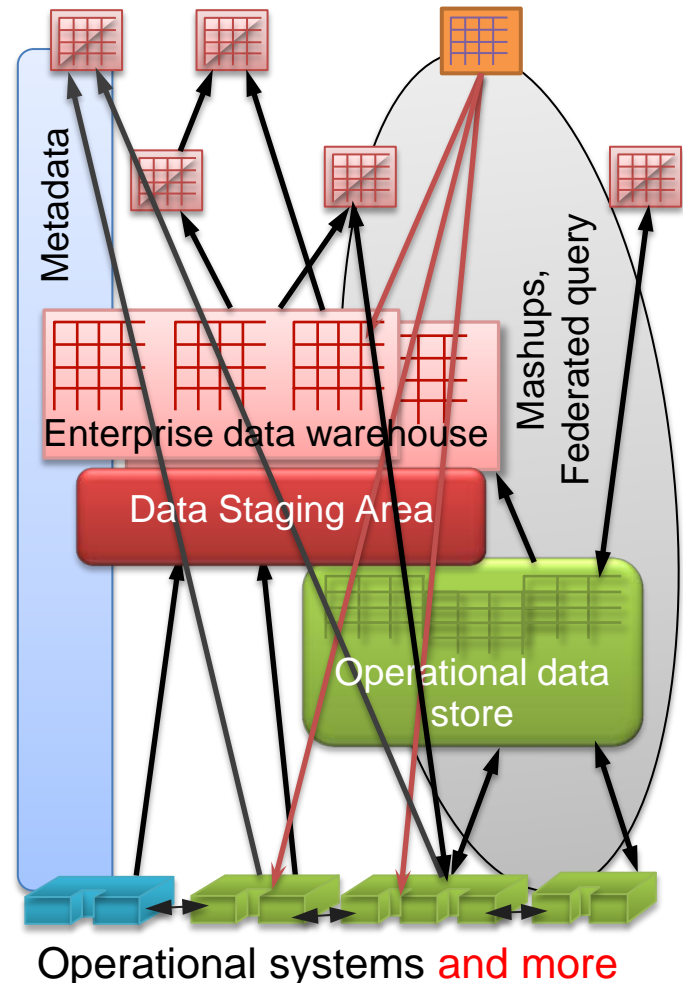
Five modern postulates for highly evolved business.

1 *Modern business processes seamlessly combine action-taking and decision-making, and require an integrated continuum of consistent information.*

3 *The business information resource is best maintained as a single copy of each data item, with only the most minimal resort to transient layers or copies of specific subsets of data for specialized needs.*

2 *The new information architecture must be based on a comprehensive enterprise information model, spanning all types of information used in the business.*

4 *An integrated, model-based and closed-loop process environment is needed to create, maintain and use both the business information and activities.*

5 *An integrated, flexible and role-based user interface provides access to the entire business information.*

A new architecture Business Integrated Insight (BI^2)… covering all information and process
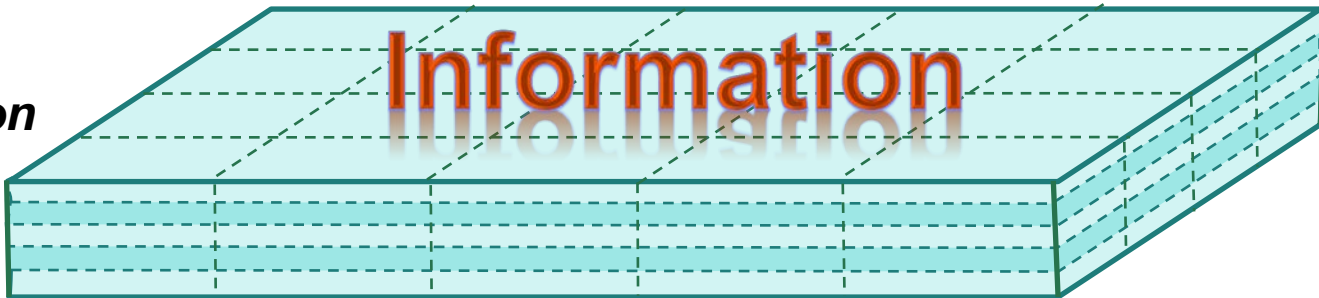
Personal Action Domain

People

Business Function Assembly

Process

Business Information Resource

Information

Enterprise-wide

See: Devlin, B. *"Business Integrated Insight (BI^2): Reinventing enterprise information management"*, (2009), http://bit.ly/BI2_White_Paper
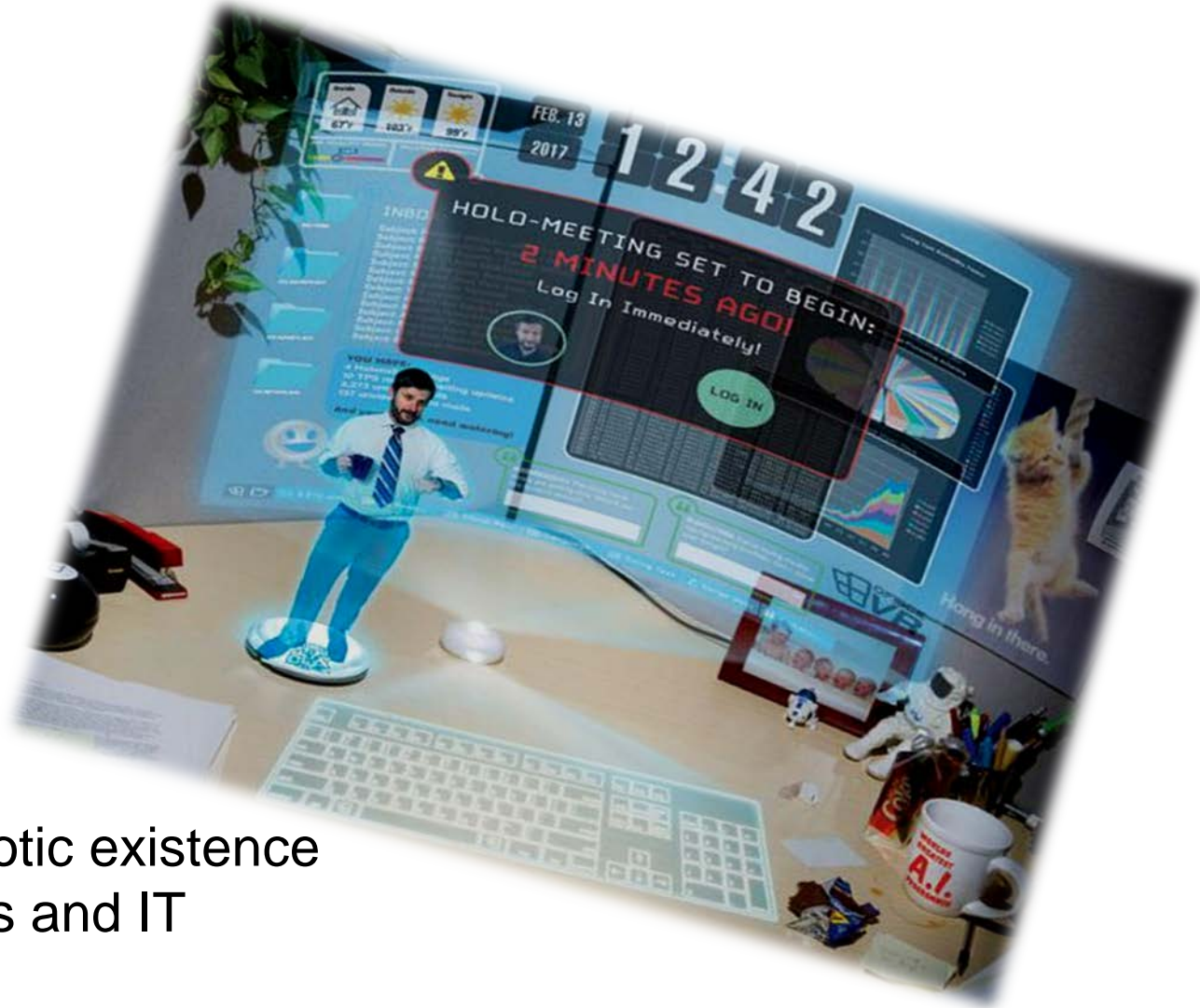
Fast-forward to 2012 –

What has changed?

Introducing the biz-tech ecosystem…

…the fully symbiotic existence
  of business and IT

The fully symbiotic existence of business and IT …

1. ## Interdependence
   - New technology enables new business possibilities;
     new business opportunities drive technology advances

2. ## Reintegration
   - Silos in business and IT are obvious to Web-savvy customers;
     coherence becomes mandatory

3. ## Cross-over
   - Business people need IT skills to see how to recreate the business
     with new technology;
     IT people need business acumen to see how to satisfy business
     needs in new ways with emerging technology

Business and IT

...no more

Beauty and the Beast

Biz-tech ecosystem – example 1

- **Business Intelligence reinvents Retail**

    – Starting with till scanning systems…

    – To the warehouse and…

    – All the way back to the manufacturer

Biz-tech ecosystem – example 2

- **The Web recreates the library**
  - The 3 Rs – recording and researching reality
  - Wikipedia – over 20 million articles

- **Democracy replaces authority**
  - Who vouches for truth?
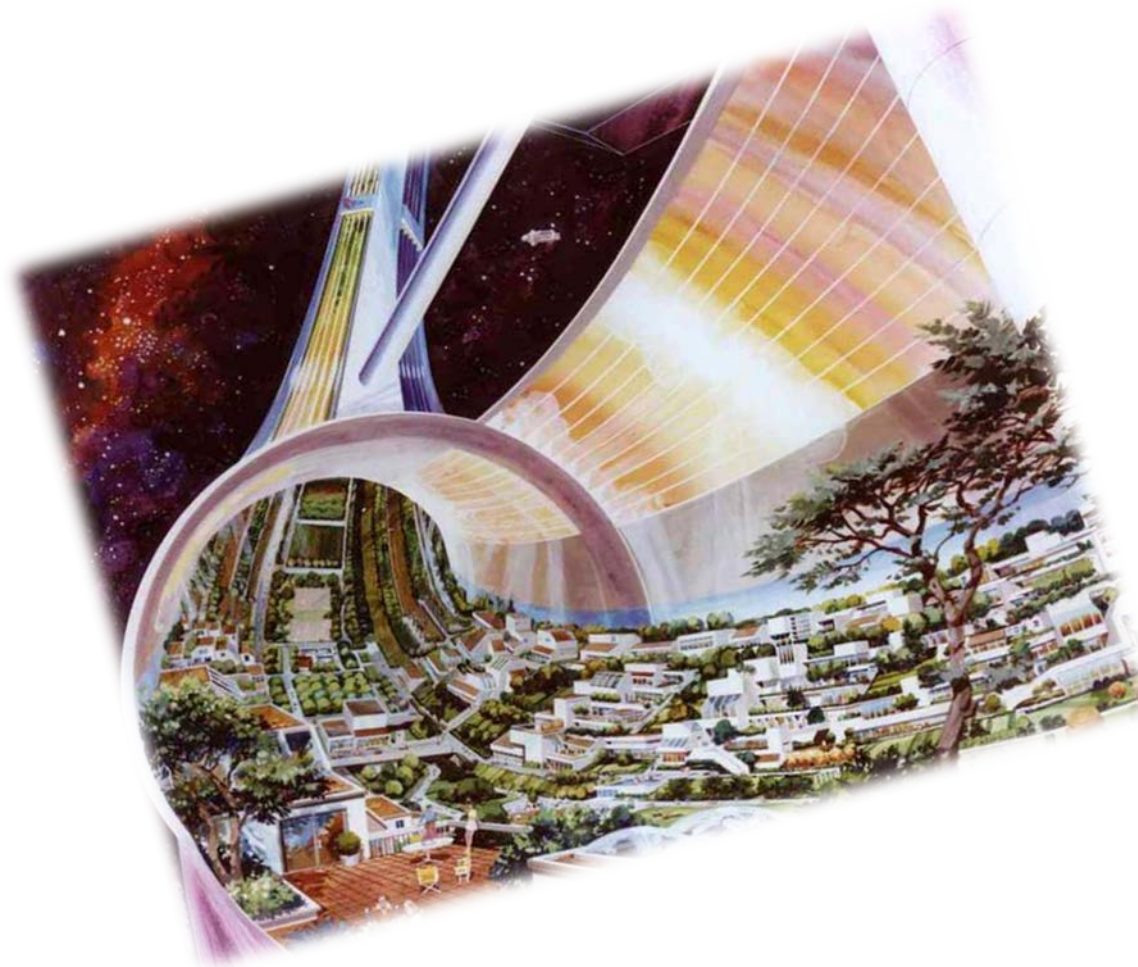 - Whither copyright and intellectual property?

Biz-tech ecosystem – example 3

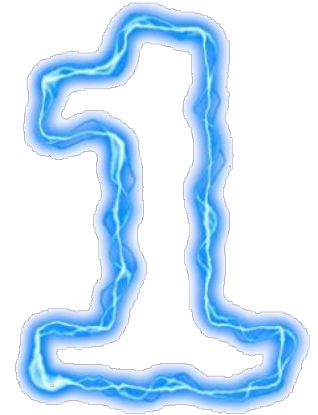- Big data redefines automobile insurance

    - Pay as your drive

    - Spreading risk becomes avoiding risk

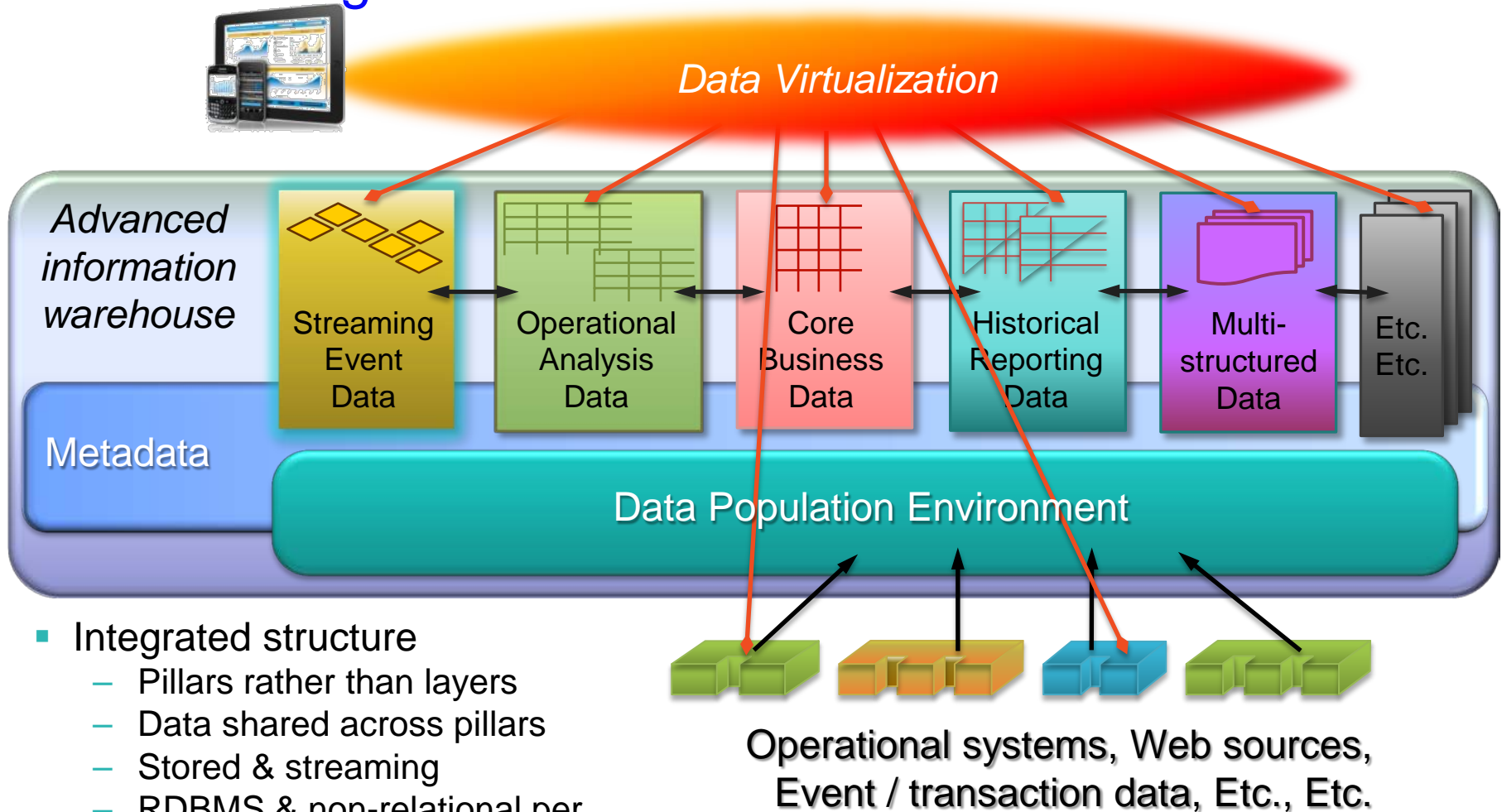# Three new architectural pictures for 2013...

Architecture of informational systems

- BI^2 is still valid, but must be approached in stages

- For 2013, we need to remove or reduce the layering in our informational systems

Introducing the advanced information warehouse

Data Virtualization

Advanced information warehouse

| Streaming Event Data | Operational Analysis Data | Core Business Data | Historical Reporting Data | Multi-structured Data | Etc. Etc. |

Metadata

Data Population Environment

- Integrated structure
  - Pillars rather than layers
  - Data shared across pillars
  - Stored & streaming
  - RDBMS & non-relational per processing needs
  - Metadata (& models) shared across pillars

Operational systems, Web sources, Event / transaction data, Etc., Etc.

Dealing with new information types
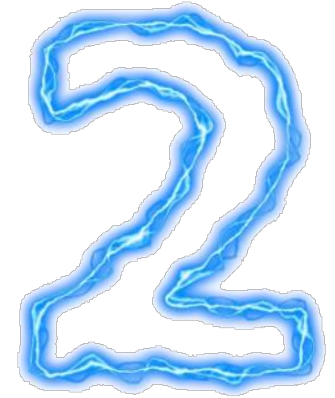
- Big data poses real issues beyond volumes, velocity and variety

- Big data challenges our fundamental beliefs about the relationship between data and knowledge
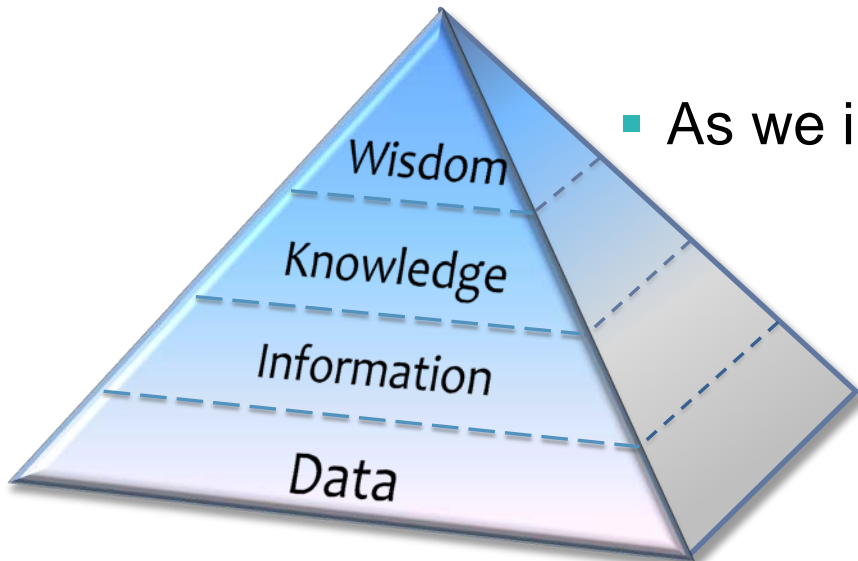

2

Data, Information, Knowledge, Wisdom ...but what about meaning?
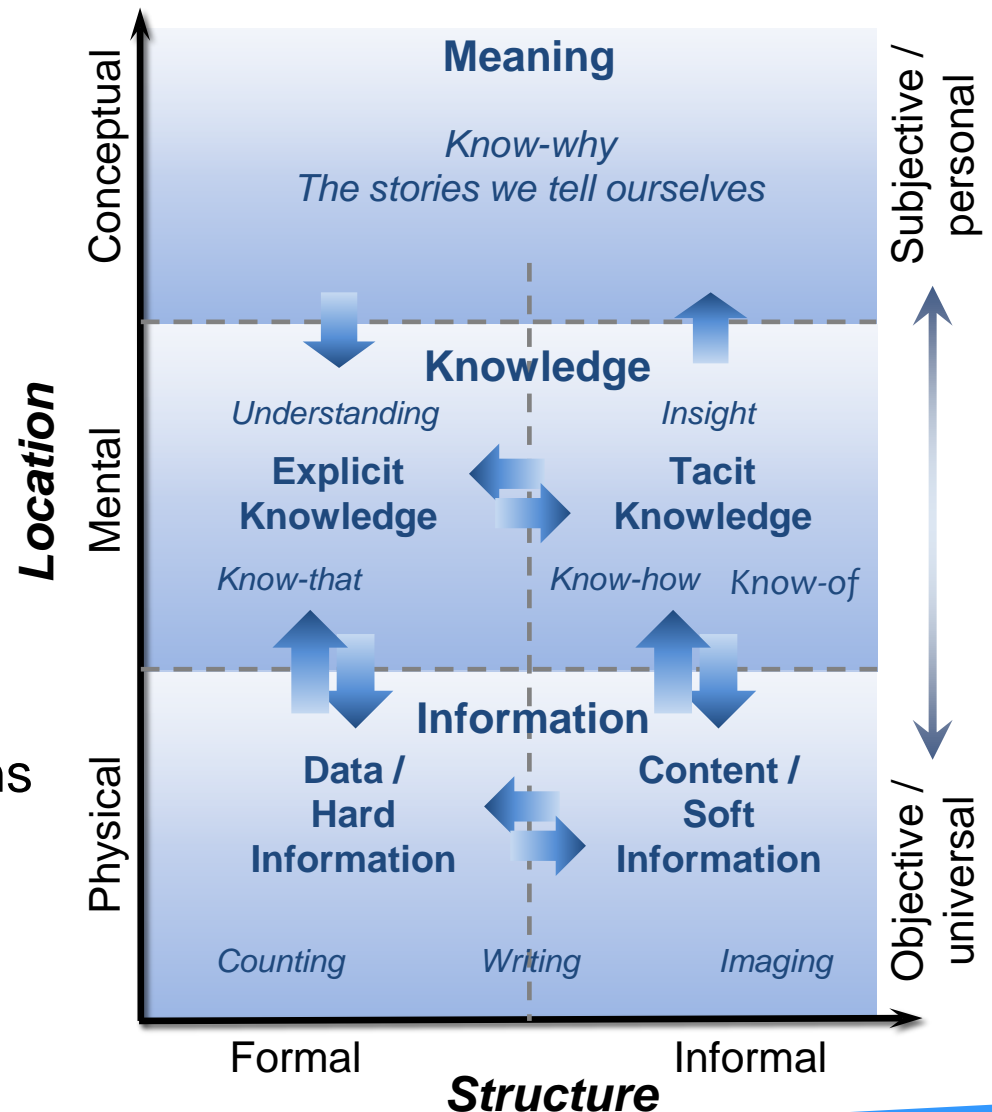
- Opening Pandora's Box

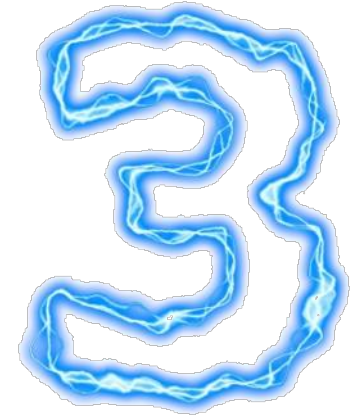- As we introduce new information types, does DIKW still make sense?

Introducing m3 – the modern meaning model

- **Information** – stored signs and symbols to describe and communicate about the world and our thoughts
  - Data (hard), Content (soft)

- **Knowledge** – the total of our experience, values, understanding & insight

- **Meaning** – the interpretations that people put on "reality", stories that differ from person to person

Meaning

Know-why
The stories we tell ourselves

Knowledge

Understanding   *Insight*

Explicit Knowledge   **Tacit Knowledge**

Know-that   *Know-how*   *Know-of*

Information

Data / Hard Information   **Content / Soft Information**

Counting   *Writing*   *Imaging*

Conceptual / Mental / Physical — *Location*

Formal   Informal — **Structure**

Subjective / personal — Objective / universal

Introducing the new decision makers

- Much of our decision making is far from rational

- Collaboration has always played a role

- For Millennials, team play is a core practice
  - (And good for innovation, too)

Decision making today – a solitary game

- Traditional BI is focused on Investigation
  - Mostly driven by a single individual, Insights reviewed at the end

Investigation

(Intention)

Integration   Interpretation

Challenge

Intuition

Formal Information

Insight
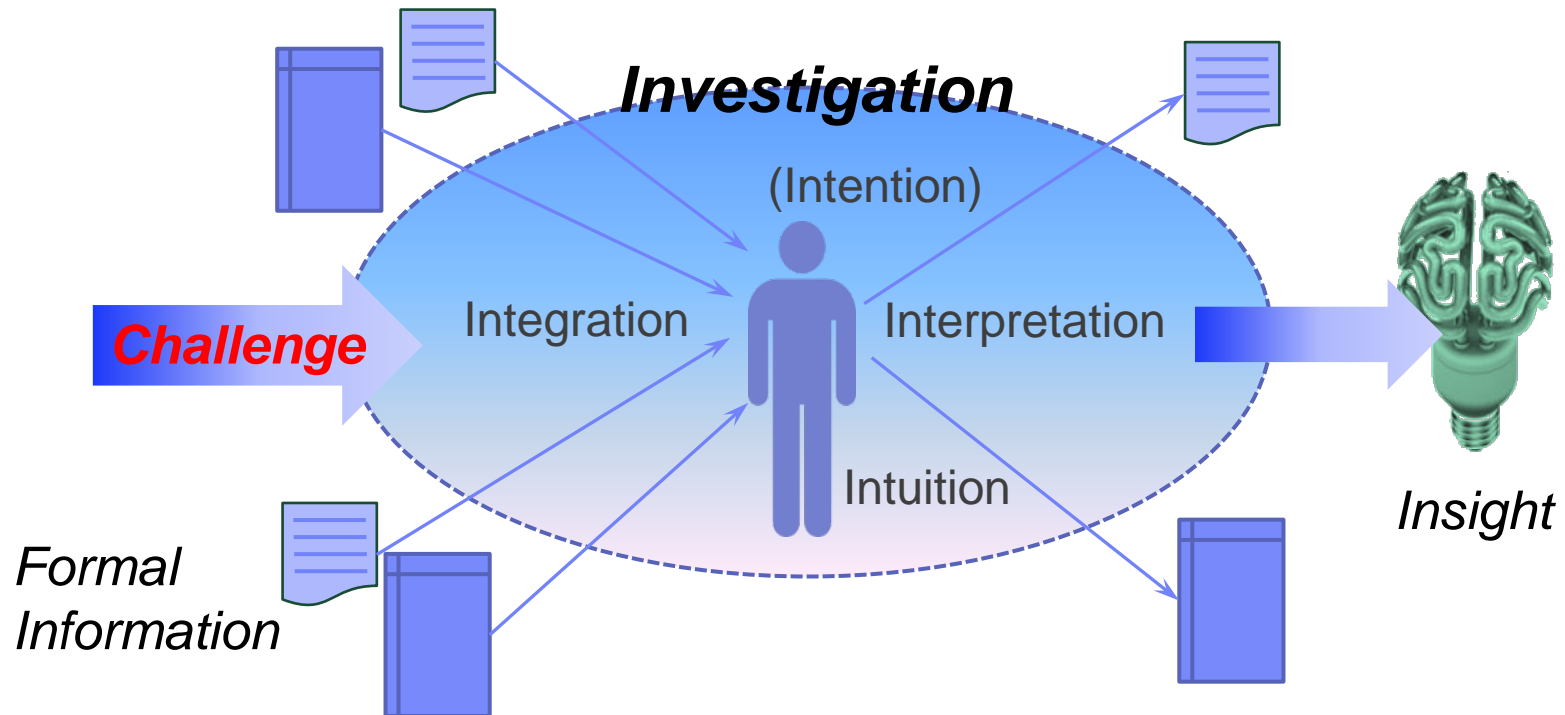
- Formal information is key driver
  - Mostly hard data

Collaboration-oriented user function:
Web 2.0 and Enterprise 2.0 driving change.

- "Collaborative BI" shows an increasing recognition that:

- All organisations are social networks as much as (or more than) hierarchies

- Social interactions and behaviour are key drivers of real and lasting innovation

- Technology is now advanced enough to support social interactions and link them to traditional business function

- Current and future employees expect Internet norms of connectivity and networking to carry into work life

Mobile computing – tablets and smart phones – are key new information targets and data sources

- **People in constant motion and in constant contact**
  - Meetings, on- / off-site, travelling

- **Tablets provide excellent form factor for information dissemination**
  - Mobility, ease of use

- **Smart phones and tablets becoming major sources of data**
  - Traditional: location, transactions, etc.
  - Emerging: e-mail, images, etc.
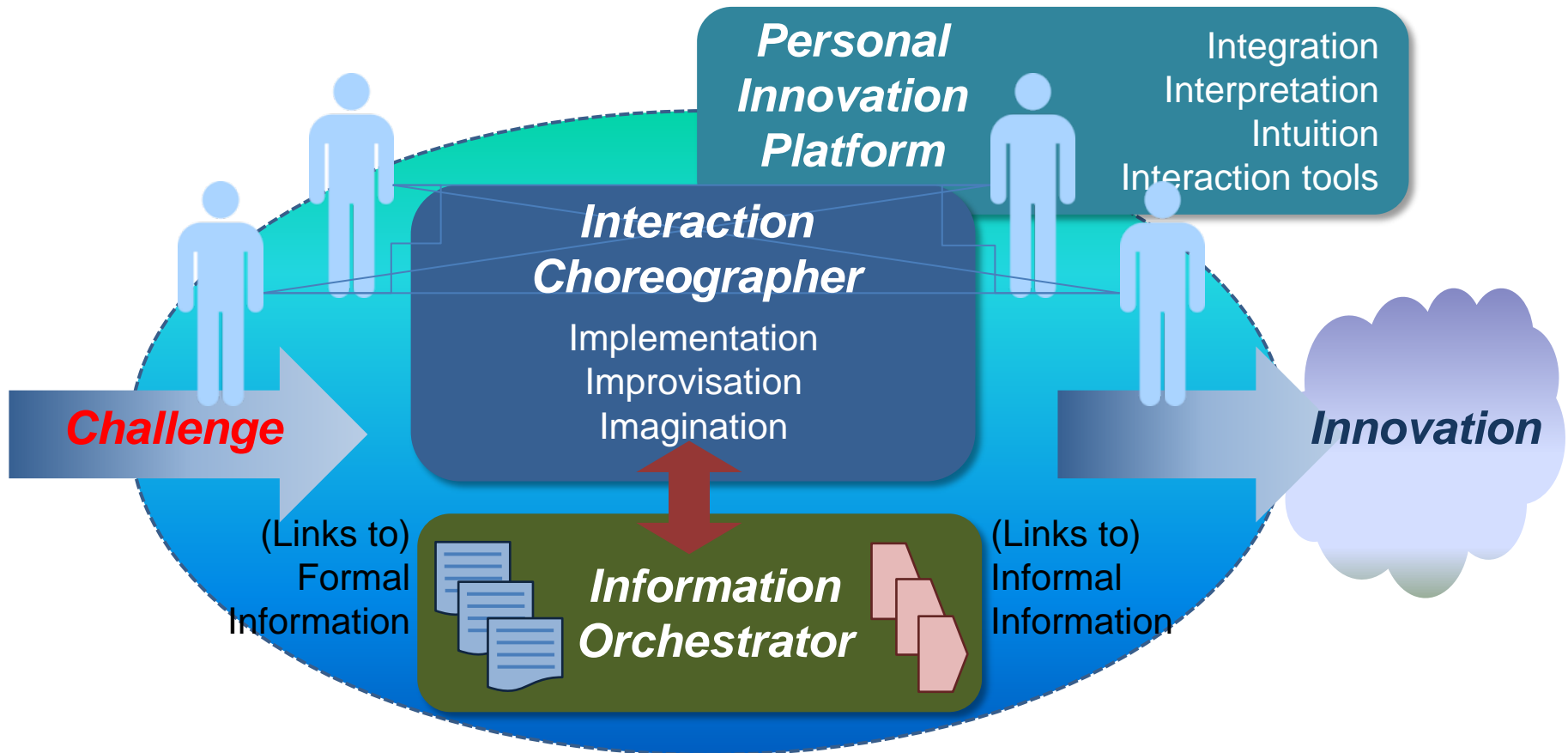 - *And recording of interactions (informal information)*

28

Introducing the iSight Model – Team-based decision making

- Decision-making teams are driven by Interaction

Interaction

Formal Information

Challenge

Implementation

Improvisation

Imagination

Informal Information

Idea
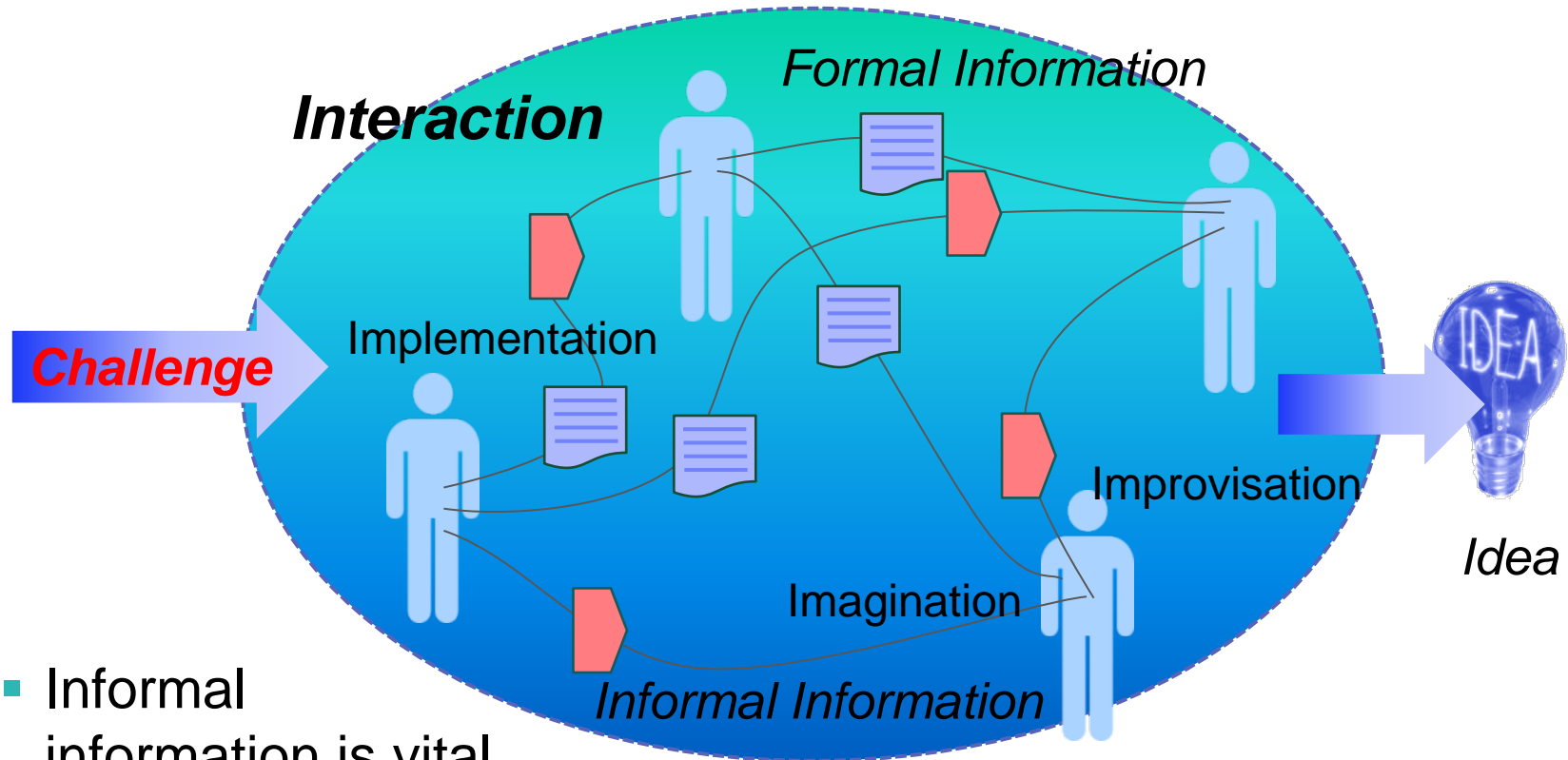
- Informal information is vital
  - Communication (digitally recorded)

iSight links Interaction and Investigation – from information to innovation

Personal Innovation Platform

Integration
Interpretation
Intuition
Interaction tools

Interaction Choreographer

Implementation
Improvisation
Imagination

Challenge

Innovation

(Links to) Formal Information

Information Orchestrator

(Links to) Informal Information

Further reading: B. Devlin, *"iSight for Innovation – Breakthrough collaboration for decision making",* (March 2012), http://bit.ly/iSights

Conclusions:
2012 – if not the end, then surely a beginning!

1. **Overall – simplify the BI environment**
   - Less layers, less copies, less ETL
   - Recognise the emerging biz-tech ecosystem

2. **Big Data – forget the hype, but do evaluate**
   - Business opportunities may exist in unexpected places
   - Recall that big data has very different characteristics

3. **Enable innovation through team working**
   - Collaborative decisioning vs. collaborative BI
   - The emerging role of informal information

One final thought...

ACM: The Learning Continues

- Questions about this webinar? learning@acm.org

- Business Intelligence/Data Management Tech Pack: http://techpack.acm.org/bi/

- ACM Learning Center: http://learning.acm.org

- ACM SIGKDD: http://www.sigkdd.org

Association for Computing Machinery

Advancing Computing as a Science & Profession

33