Patrick Derde

Contact information:

Email: p.derde@bizzdesign.com
      patrick.derde@envizion.eu
Mobile: +32 (0)497 302387
Web:    www.bizzdesign.com
        www.envizion.eu

Patrick is a TOGAF 9 and ArchiMate 2.1 certified enterprise architect. He has more than 20 years experience in the different domains of enterprise architecture especially in financial institutions. He has a master in Commercial Engineering at the university KU Leuven and he also graduated in actuarial sciences. In 2009 he was awarded and nominated as a fellow of the KU Leuven Hogenheuvel Institute. He is founding member and director of the Data Management Association (www.dama-belux.org). He is also managing partner of BiZZdesign Belgium (www.bizzdesign.com) and Managing Partner of Envizion (www.envizion.eu).
Mission BiZZdesign

• BiZZdesign helps building strong organizations by providing them the agility needed to act upon today’s and tomorrow’s challenges

• We offer an integral approach to designing, improving and managing businesses effectively

• The integral approach consists of
  – proven and easy to use software tools,
  – best practice models and methods,
  – training,
  – business consultancy
BiZZdesign offers a complete solution!

<table>
<thead>
<tr>
<th>Tooling</th>
<th>Consultancy</th>
<th>Best practices</th>
<th>Training</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>BMI</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>► Business model canvas</td>
<td>► Strategy development</td>
<td>► Business Model Foundation</td>
</tr>
<tr>
<td><strong>EA</strong></td>
<td>► Architect</td>
<td>► Architecture consultancy</td>
<td>► Archimate 2.0 &amp; Architect</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>► TOGAF 9.1 Level1&amp;2</td>
</tr>
<tr>
<td><strong>BPM</strong></td>
<td>► BiZZdesigner</td>
<td>► Business/ Lean consultancy</td>
<td>► BiZZdesigner foundation</td>
</tr>
<tr>
<td></td>
<td>► TDM</td>
<td>► Design/models</td>
<td>► Lean Green Belt</td>
</tr>
<tr>
<td></td>
<td>► Lean Coach</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>GRC</strong></td>
<td>► GripManager</td>
<td>► Management and/or project</td>
<td>► Risk Mngt.</td>
</tr>
<tr>
<td></td>
<td>► RiskManager</td>
<td>management</td>
<td>► Requirements Mngt.</td>
</tr>
</tbody>
</table>
Envizion

Vision

Information is a corporate asset

Mission

Support the design and implementation of strategic, aligned, accurate and trusted business information. Allowing to capitalize on the information assets and to share information throughout the company and with the stakeholders, resulting in an efficient and effective organization.
Envizion offers Information Architecture Capabilities

---

**Client Business Capabilities**

- Corporate Governance
- Business Strategy
- Business Architecture

**Envizion Information Capabilities**

- EA Governance
- EA Strategy
- EA Architecture
- Information Management
- Information Development
- Meta Data Management
- Information Quality Management
- Information Security Management
- Data Warehousing & Business Intelligence Management
- Reference & Master Data Management

**Client Technology Capabilities**

- IT Governance
- IT Strategy
- IT Architecture
- IT Operations

---

**Envizion Business Capabilities**

- Business Operations

---

**Envizion Technology Capabilities**

- IT Governance
- IT Strategy
- IT Architecture
- IT Operations
Big Data – Use Cases and Architecture
Agenda:

- What is Big Data?
- Why Big Data?
- Big Data Technology Architecture
- Big Data Application Architecture
- Big Data Business Architecture
- Big Data Business Use Cases
Big Data

Information, data and drawings embodied in this document are strictly confidential and are supplied on the understanding that they will be held confidentially and not disclosed to third parties without the prior written consent of Envizion CVBA.

www.tweetping.net

Big Data Assessment

more history

more detailed data

More tables

More columns

Data

comes from

goes to

Data Channels

Big Data

Social media

Multi Channel

Sensors

Internet of Things

goes to

comes from

Location

goes to

comes from

THINGS

Everywhere
What is Big Data?

• A general used term to refer to datasets that are characterised by

  – **Volume**: high amount of data
  – **Velocity**: high speed data is coming in
  – **Variability**: variance in meaning and structure (depending on the context)
  – **Variety**: of data formats (structured, unstructured, semi structured)
  – **Veracity**: uncertainty of meaning and quality, corresponding to reality
Big Data Concerns: Technical Input Concerns

• Volume of
  – data to be processed
  – Things sending data
  – Transactions to be processed

• Velocity of data
  – coming in
  – outdated

• Variety of
  – data structures for data of the same class (Structured, Unstructured, Semi structured)
  – Devices/Interfaces through which data is passed
  – Data Stores

• Variability of
  – data classes

• Veracity of data, uncertainty of
  – Meaning
  – quality
Big Data: Technical Process Concerns

- High Volume Transactions
  - Input
  - Process
  - Output
- Complex Logic
  - Data Search
  - Data Integration (data scattered over different locations)
  - Data (Artificial) Intelligence
- High Speed
  - Near real time processing (zero latency)
  - Data access
  - Knowledge delivery

- Operational BI
  - Real time BI
  - Real time Analytics
  - Operational Intelligence
  - Operational Analytics
  - Active Data Warehousing
Agenda:

• What is Big Data?
• Why Big Data?
• Big Data Technology Architecture
• Big Data Application Architecture
• Big Data Business Architecture
• Big Data Business Use Cases
Why Big Data? To create Value

Big Data ≠ Value

- Big Data
- Internet of Things
- Social media
- Multi Channel
- Sensors
- Everywhere

Information, data and drawings embodied in this document are strictly confidential and are supplied on the understanding that they will be held confidentially and not disclosed to third parties without the prior written consent of Envizion CVBA.
Why Big Data?

Business wants to use data to create knowledge for effective and efficient decision making.
Big Data: IT Goals

- Lower Total Cost of Ownership
  - Lower License costs
  - Less Time spend on physical Database Design
  - Less Time spend on database administration

- Faster Query Processing
  - Sub-second Response Times
  - More users
  - Different groups of users
  - More queries
  - More complex queries
  - More queries on detailed data

- Increased Flexibility
  - More tables and columns
  - Change existing structures
  - Changing Partitioning schema’s
  - Changing buffer spaces
  - Changing aggregated tables
  - Changing Aggregated columns
  - Changing ...
Big Data: Technology Requirements

Information, data and drawings embodied in this document are strictly confidential and are supplied on the understanding that they will be held confidentially and not disclosed to third parties without the prior written consent of Envizion CVBA.
Agenda:

- What is Big Data?
- Why Big Data?
- Big Data Technology Architecture
- Big Data Application Architecture
- Big Data Business Architecture
- Big Data Business Use Cases
What is an Architecture?

- An Architecture is the fundamental organization of a system, embodied in:
  - its components,
  - their relationships to each other and the environment,
  - and the principles governing its design and evolution.

Adapted from ANSI/IEEE Standard 1471-2000
## ArchiMate - Architecture Components

<table>
<thead>
<tr>
<th>Passive structure</th>
<th>Behavior</th>
<th>Active Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Business product</td>
<td>Business service</td>
<td>Business interface</td>
</tr>
<tr>
<td>Contract</td>
<td>Business event</td>
<td>Business role</td>
</tr>
<tr>
<td>Business object</td>
<td>Business function</td>
<td>Business actor</td>
</tr>
<tr>
<td>Data object</td>
<td>Application service</td>
<td>Application interface</td>
</tr>
<tr>
<td>Artifacts</td>
<td>Application function</td>
<td>Application component</td>
</tr>
<tr>
<td></td>
<td>Infrastructure service</td>
<td>Infrastructure interface</td>
</tr>
</tbody>
</table>

**Main concepts**

- Business layer
- Application layer
- Technology layer
What is Big Data Technology Architecture?

- A Big Data Architecture is the fundamental organization of a Big Data Technical System, embodied in:
  - its Technology components,
  - their relationships to each other and the environment,
  - and the principles governing its design and evolution.

Adapted from ANSI/IEEE Standard 1471-2000
Big Data: Technology Requirements realised by Technology Architecture

Flexible Data Model

Complex Data Bases
Structured Data
Web Data
Semi Structured Data
Unstructured Data
Multiple Languages
Text
Pictures
Video
Sound

Big Data Capability
Data Mining Capability
Programs

NoSQL Database Server

NoSQL
Big Data: Technology Architecture

- Will realise Big Data Technology Requirements
- This is what is focussed on in 2013 and before
  - Bringing together Legacy Infrastructure with Big Data Infrastructure
    - NoSQL Database Servers
    - SQL Database Servers
Big Data: Graph Technology Solution Architecture

- InfiniteGraph
- AllegroGraph
- RDFStore
- Neo4j
- Vertexdb

Diagram showing technology goals and solutions:
- Lower Total Cost of Ownership
- Lower License costs
- Less Time spend on database administration
- Technology Goals
- Increased Flexibility
- Less Time spend on physical Database Design
- Faster Query Processing
- Neo4j
  - NoSQL database
  - Cypher Search Language
- More complex queries
- Distributed Databases
- Graph Data Store
- IBM IMS
- IDMS
- CDS
Big Data Architecture Components

<table>
<thead>
<tr>
<th>Passive structure</th>
<th>Behavior</th>
<th>Active Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Business layer</td>
<td>Application</td>
<td>Technology layer</td>
</tr>
</tbody>
</table>

Data Infrastructure Engineers

Neo4j
NoSQL database

Cypher Search Language
Market of NoSQL products
Agenda:

• What is Big Data?
• Why Big Data?
• Big Data Technology Architecture
  • Big Data Application Architecture
• Big Data Business Architecture
• Big Data Business Use Cases
Why Big Data? Take create Value!

Information, data and drawings embodied in this document are strictly confidential and are supplied on the understanding that they will be held confidentially and not disclosed to third parties without the prior written consent of Envizion CVBA.
Data Rich ... Information Poor ...

- Transform Data into Information and Insight
Business Intelligence

A set of *capabilities* that transform data into useful information and insight for strategic, tactical and operational decision making.

- We need to understand the data, the meaning (Semantics)
- We need common language
- We need standards
- We need data manipulation functions
- We need processing power

• Smart and Intelligent Data Usage

• **Business Value** = $f(\text{Big Data}, \text{Enterprise Data})$

We need to be able to express “What is business Value ?” in order to know what capabilities can transform data into Business Value
Big Data is about Business

Business Value = \( f(\text{Big Data, Enterprise Data}) \)

Information, data and drawings embodied in this document are strictly confidential and are supplied on the understanding that they will be held confidentially and not disclosed to third parties without the prior written consent of Envizion CVBA.
Business Value = $f(\text{Big Data}, \text{Enterprise Data})$

What are the functions $f$ to transform Data into Value?

- Semantics Engine
- Artificial Intelligence Engine
  - Self learning Engine
  - Scalable machine learning
  - Evolutionary computing algorithms
- Genetic programming
- Rule Engines
- Cognitive Computing
- Mathematics of taste
- Predictive Model Engine
- ...

Information, data and drawings embodied in this document are strictly confidential and are supplied on the understanding that they will be held confidentially and not disclosed to third parties without the prior written consent of Envizion CVBA.
Why Big Data? To create Value!

Big Data

Social media

Multi Channel

Sensors

Internet of Things

Everywhere

THINGS

Location

Data Channels

comes from

goes to

comes from

goes to

Big Data

Modern Data Scientist

Math & Statistics
- Machine learning
- Statistical modeling
- Experiment design
- Bayesian analysis
- Supervised learning: decision trees, random forests, logistic regression
- Unsupervised learning: clustering, dimensionality reduction
- Optimization: gradient descent and variants

Programming & Database
- Database systems fundamentals
- Sampling language e.g. Python
- Statistical sampling packages e.g. R
- Databases SQL and NoSQL
- Numerical algebra
- Parallel databases and parallel query processing
- MapReduce concepts
- Hadoop and HBase
- Custom reducers
- Experience with tools like AWS

Domain Knowledge & Soft Skills
- Passionate about the business
- Committed to data
- Influence without authority
- Analytical mindset
- Problem solver
- Strategic, creative, innovative and collaborator

Communication & Visualization
- Ability to engage with senior stakeholders
- Story telling skills
- Translate data-driven insights into decisions and actions
- Visual art design
- R packages like ggplot or lattice
- Knowledge of any visualization tools like Tableau, PowerBI, Qlik

Information, data and drawings embodied in this document are strictly confidential and are supplied on the understanding that they will be held confidentially and not disclosed to third parties without the prior written consent of Envizion CVBA.
Big Data: Application Architecture

- Will realise Big Data Business Requirements
- This is what is focussed on in 2014 and will lead to value creation in the future.
### Big Data Architecture Components

<table>
<thead>
<tr>
<th>Passive Structure</th>
<th>Behavior</th>
<th>Active Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data Stewards</strong></td>
<td><strong>BI Services</strong></td>
<td><strong>Subject Matter Experts</strong></td>
</tr>
<tr>
<td>Data Infrastructure Engineers</td>
<td><strong>BI Interface</strong></td>
<td><strong>Subject Matter Scientists</strong></td>
</tr>
<tr>
<td>Data:</td>
<td><strong>BI Capability</strong></td>
<td><strong>Neo4j NoSQL database</strong></td>
</tr>
<tr>
<td>- Enterprise Data</td>
<td>- Semantic Engine</td>
<td>- Cypher Search Language</td>
</tr>
<tr>
<td>- External Data</td>
<td>- Intelligent Mining Function</td>
<td></td>
</tr>
<tr>
<td><strong>Data Stewards</strong></td>
<td>- Artifical Intelligence Engine</td>
<td></td>
</tr>
<tr>
<td><strong>Data Infrastructure Engineers</strong></td>
<td>- Biostatistical Learning</td>
<td></td>
</tr>
<tr>
<td><strong>Data Infrastructure Engineers</strong></td>
<td>- Scientific Domain Functions</td>
<td></td>
</tr>
<tr>
<td><strong>Data Infrastructure Engineers</strong></td>
<td>- Evolutionary Computing Engineering</td>
<td></td>
</tr>
</tbody>
</table>

**Business Layer**

**Application Layer**

**Technology Layer**
10 Big Data Companies by Revenue

1. IBM
2. Hewlett Packard
3. Teradata
4. Dell
5. Oracle
6. SAP
7. EMC
8. Cisco
9. PwC
10. Microsoft

http://wikibon.org/wiki/v/Big_Data_VENDOR_Revenue_and_Market_Forecast_2012-2017#Big_Data_VENDOR_Revenue
Agenda:

• What is Big Data?
• Why Big Data?
• Big Data Technology Architecture
• Big Data Application Architecture
• **Big Data Business Architecture**
• Big Data Business Use Cases
Big Data Business Architecture Components

Passive structure

Behavior

Active Structure

Business layer

Application layer

Technology layer
Use Cases 2015 – Beyond the sand box?

Business Value = $f(\text{Big Data}, \text{Enterprise Data})$

• Marketing
  – 360 view on the customer, digital shopping
  – Optimize sales funnel
• Logistics/Transport
  – Physical Delivery, just enough, just in time
• Supply Chain
  – Predict how much of what, when and where
• Telecom
• Research Institutions
• Financial Institutions
  – Risk Management
  – Stock market component evolution
• Government
  – Detecting costly tax-evasion schemes
  – Civil Security
  – Traffic control
Use Cases 2015 – Beyond the sand box?
Business Value = $f(\text{Big Data, Enterprise Data})$

- **Energy**
  - Laying out Wind farms
  - Smart meters
  - Predicting Power usage
- **Health care**
  - Studying and categorizing the beats in blood pressure data in order to predict drops and spikes
  - Computer based diagnosis
  - Pharmarmacy, development of medicines
- **Lawyers**
  - Text Analysis
  - International laws
- **Weather Forecasts**
- ...
Telecom : Data Clearing
## Near real time analytics

<table>
<thead>
<tr>
<th>Company</th>
<th>Industry: Telecom</th>
</tr>
</thead>
<tbody>
<tr>
<td>MACH</td>
<td>Use Case Type: Operational Analytics</td>
</tr>
</tbody>
</table>

**Purpose:**

Mach is data clearing company.

**Problem:**

Network providers pay each other for the use of each others network when customers are roaming across the borders of the network. Because of the huge competition and regulation telephone customers agree a limited amount of data usage. Because of the delay between the data-usage and the processing of the usage in the transaction systems, telecom providers loose a lot of money because they are not allowed to bill the excess usage to the customer.

**Solution:**

Near real time exchange of the data usage.

Optimizing enterprise application and business process architecture
Flanders Research Information Space

"That book is still floating around the system."

Information, data and drawings embodied in this document are strictly confidential and are supplied on the understanding that they will be held confidentially and not disclosed to third parties without the prior written consent of Envizion CVBA.
Company: EWI

Industry: Government

Use Case Type: Information Exchange

Purpose:

Exchanging research results.

Problem:

Government is investing in research and development. With public money. Research Institutes request support for R&D. Local investments, worldwide research. Different languages. Different classification mechanisms. How to exchange the research output? How to evaluate the return on investment? How to derive research and development policy? How can industry benefit from the results?

Solution:

Structured Research Information Space.

Standardisation of Research Meta Data.
Marketing: 360° view, Product development

A Great Company gains market share in a downturn

Copyright Moshe Levi

“Forget about softness. Tests show we could triple sales if we add pepper.”

Information, data and drawings embodied in this document are strictly confidential and are supplied on the understanding that they will be held confidentially and not disclosed to third parties without the prior written consent of Envizion CVBA.
Use Case: Marketing
360° view on the customer (most popular)

• Value:
  – influence customer’s sales decision
  – Predict future sales based on patterns identified by means of the huge
    amount of multidimensional data gathered from high volume customer
    behaviour

• Big Data?
  • Online shopping. What are shoppers doing on their site? What pages
    visits, how long, what clicks, when leaves?
  • What’s the sentiment of the customer?
  • Customer commercial psychology
  • Customer group psychology

• Business Intelligence?
  – Predict Buy Pattern, behavior analytics
  – Online real time “help to buy” services
# Investment Decisions

<table>
<thead>
<tr>
<th>Company</th>
<th>Industry: Finance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Use Case Type: Predictive Analytics</td>
</tr>
</tbody>
</table>

## Purpose:

Using tweet mining for investment clues.

Based on sentiment analysis in the tweets …
## Selling Data

<table>
<thead>
<tr>
<th>Company</th>
<th>Industry: Finance</th>
</tr>
</thead>
<tbody>
<tr>
<td>MasterCard</td>
<td>Use Case Type: Behavioral Analytics</td>
</tr>
</tbody>
</table>

**Purpose:**

With 1.8 billion customers, MasterCard is in the unique position of being able to analyze the behavior of customers in not only their own stores, but also thousands of other retailers. The company teamed up with Mu Sigma to collect and analyze data on shoppers’ behavior, and provide the insights it finds to other retailers in benchmarking reports.
Predict Security Threats

<table>
<thead>
<tr>
<th>Company</th>
<th>Industry: Finance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Use Case Type</td>
<td>Behavioral Analytics</td>
</tr>
</tbody>
</table>

**Purpose:**
Rabobank analysed criminal activities at ATMs to determine factors that increased the risk of becoming victimized. It discovered that proximity to highways, weather conditions and the season all affect the risk of a security threat.
Wind Energy Farm design

Information, data and drawings embodied in this document are strictly confidential and are supplied on the understanding that they will be held confidentially and not disclosed to third parties without the prior written consent of Envizion CVBA.
Use Case: Energy
Wind Energy Farm Design

• Problem:
  – Where to position the Wind Turbines given a very high dimensionality
  – Hundreds of turbines are involved

• Big Data?
  – Complex Problem:
    • Characteristics of the turbine – energy gain, height, proximity cone
    • How much wind is required for the site?
    • Where does the wind comes from and in what quantities?
    • The topographical conditions of the land
    • The way the wind sweeps through it
    • The wake effect of one turbine on the turbines behind it
      – Flow modeling, to calculate the loss behind each turbine
      – Parallelized evolutionary algorithms

• Value?
  – Efficiency gains in optimizing the placement of turbines
  – Ability to scale up for wind farms on a bigger scale than ever thought was possible
Health Care: making sense of clinical data
Use Case: Health Care
Making sense of clinical data

• Problem:
  – How to derive medical insight from the growing volume of physiological data collected from body sensors and predict disease to avoid effective illness?
  – Data includes everything from blood pressure, heart beat, ECG, sleep patterns, ...

• Big Data?
  – Complex Problem:
    • Many dimensions, high diversity of parameters
    • High volume of Signal level information
    • Characteristics of every heart beat and wave form
    • Historical segments
  – Parallelized evolutionary algorithms

• Value?
  – Predict individual medical issues based on patterns identified by means of the huge amount of multi dimensional data analysed
  – Self service diagnoses ... The profession of a doctor will be under pressure
Lawyers ... Semantic Intelligence... Decision theory

Legal Advice: $5
Accurate Legal Advice: $5,000

Helping You
Help Yourself

THINGS

comes from

goes to

comes from

goes to

Information, data and drawings embodied in this document are strictly confidential and are supplied on the understanding that they will be held confidentially and not disclosed to third parties without the prior written consent of Envision CVBA.
Closing Remarks

• What’s your business case for Big Data?
  – Cost reduction?
  – Solving New problems?
  – Innovation?

• Are you sure you need it?
  – Modeling and engineering is still needed
  – Unstructured <> no need to know the meaning
  – Combine Structured and Unstructured

• Data Management and Data Governance will not be replaces by Big Data

• A solution today is not the best solution tomorrow