



S5: Interplay of data and knowledge in system analysis
TNO-ESI Symposium

Robert Jan van Wijk, ASML D&E
Eindhoven, 2017.10.03

The rise of Big Data

“90% of the world’s data has been generated in just the past two years”

The rise of Big Data

“90% of the world’s data has been generated in just the past two years”

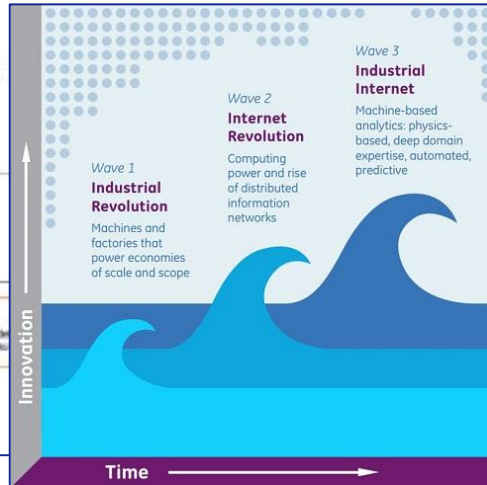
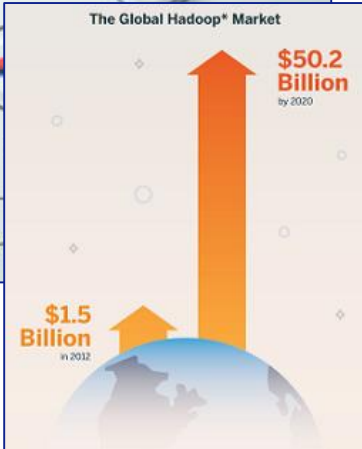
IMPACT of Big Data Play In Your Industry?



Why **61%** of oil and gas executives believe big data and analytics will be critical over the next two years

sectors are positioned for greater gains from use of big data

- Cluster A
 - Cluster B
 - Cluster C
 - Cluster D
 - Cluster E
- Bubble sizes denote relative sizes of GDP



This McKinsey Report says How Big Data Will Impact Different

Industries, But Whatever be Your Industry, you can't survive Without Big Data after the Next 10-15 Years!

So be the Early Bird.

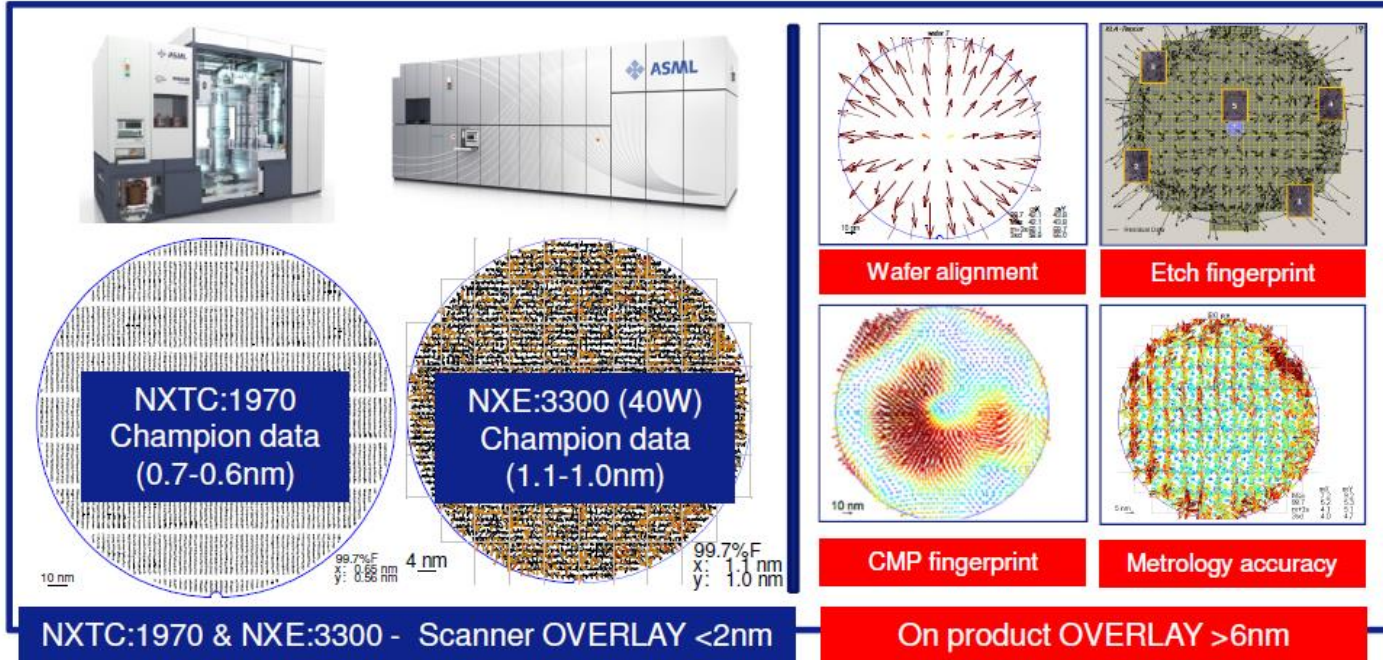


Example from industry



- An ASML lithographic scanner measures and stores 10,000 variables per silicon wafer.
- And that's just one part of the chip manufacturing process.

Example from industry



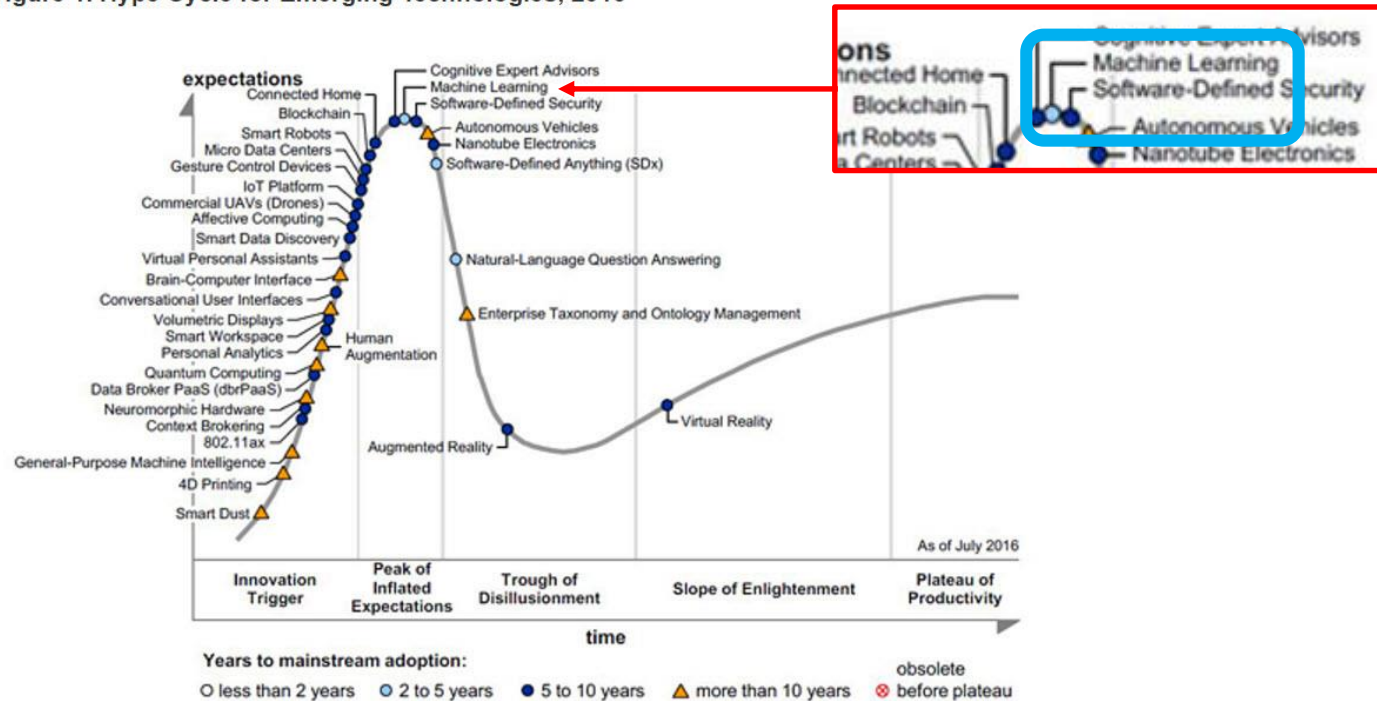
LithoInsight enables the customer to optimize their entire production process. “Holistic approach”.

Turning Big Data into Effective Information

- “Big” data and multivariate relationships cannot be handled by a human analyst/domain expert.
- Data science/machine learning provide powerful and efficient methods to demonstrate relationships, without physical understanding.
- Problem solved?

Big Data or a Big Bubble?

Figure 1. Hype Cycle for Emerging Technologies, 2016



Source: Gartner (July 2016)

Source: Gartner (August 2016)

Some thoughts on industrialization

Getting a working model is only 10% of the work.

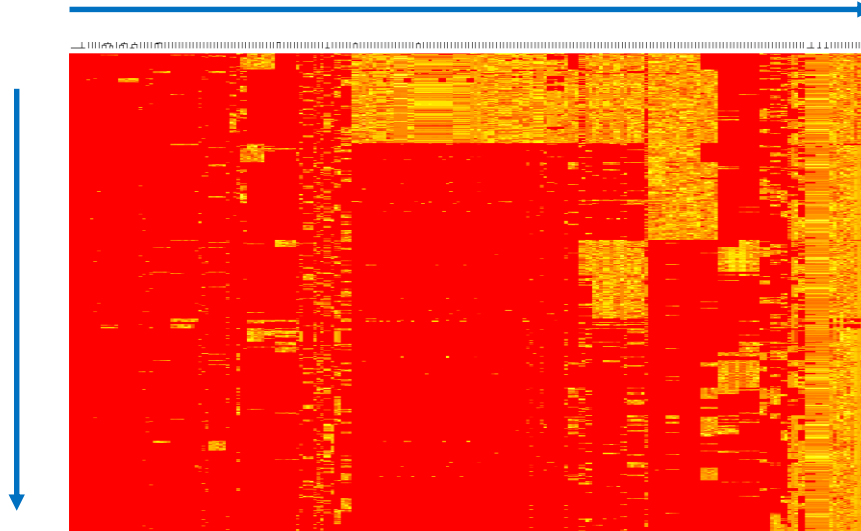
Industrial use requires:

- Scalability (from lab training set to full high-volume)
- Maintainability (flexible adaptation to process changes without total retraining)
- Robustness (insensitivity to environmental changes)
- Coping with missing data

Example missing data

150 product and process variables

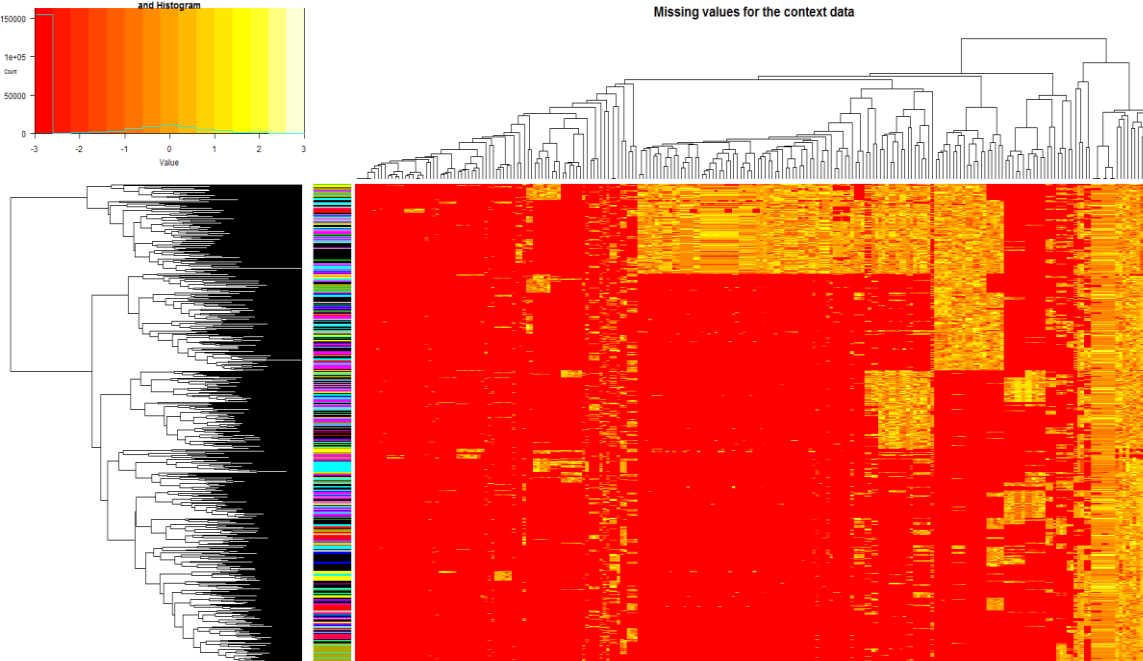
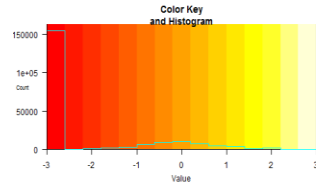
900 silicon wafers
(35 lots, 9 different chip designs)



Missing values
Measured values

Missing data: 75%

Example missing data



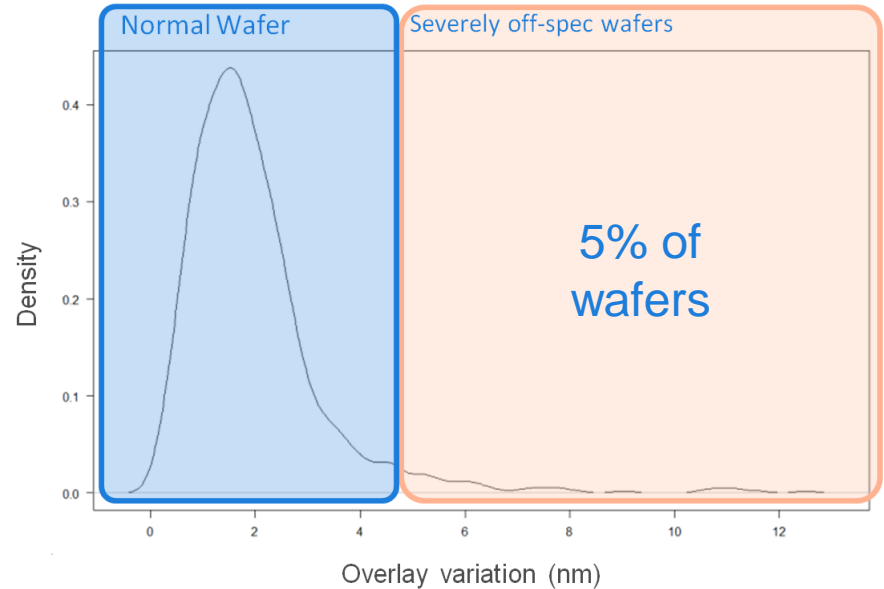
- Missing values
- Measured values

Blind classification sees missing data as the most prominent feature...

Some thoughts on industrialization

Big data, or really?

- In a well-controlled industrial process >99% of the data is within limits.
- Only outliers are potentially interesting: scarce data after all...



Some thoughts on industrialization

Getting a working model is only 10% of the work.

Industrial use requires:

- Scalability (from lab training set to full high-volume)
- Maintainability (flexible adaptation to process changes without total retraining)
- Robustness (insensitivity to environmental changes)
- Coping with missing data
- Dealing with sparseness (only exceptions bring value)
- **Combining data science with human domain expertise**

Some thoughts on industrialization

Getting a working model is only 10% of the work.

Industrial use requires:

- ➔ Scalability (from lab training set to full high-volume)
- ➔ Maintainability (flexible adaptation to process changes without total retraining)
- Robustness (insensitivity to environmental changes)
- ➔ Coping with missing data
- ➔ Dealing with sparseness (only exceptions bring value)
- ➔ **Combining data science with human domain expertise**

Presentations

S5 Interplay of data and knowledge in system analysis

- Dr. Michael Borth – TNO ESI

The Fusion of Data and Knowledge in System Models

As big data does not come to the rescue of smart systems

- Peter Mas M.Sc.Eng. – Siemens PLM

Deep Learning and Deep Analysis

The interplay between neural networks and system simulation to support large data processing

The image features the ASML logo in a bold, dark blue, sans-serif font. The logo is positioned on the left side of the frame. The background is a light blue gradient with abstract, flowing white and light blue wave-like patterns that sweep across the scene from the bottom left towards the top right. The overall aesthetic is clean, modern, and professional.

ASML