



Data Driven Healthcare – Challenges and Opportunities

Dr. Stefan Wess

CEO Empolis Information Management

Watson changed history in February 2011



Mega Trend: Smart, Social and Online

There are over 1,5 billion iOS and Android units in use around the globe today

Mobile and smart units are currently growing in numbers...

- **10 times faster** than **personal computers** in the 80s
- **twice as fast** as the amount of **Internet users** in the 90s
- **3 times faster** than the **social networks** in the 2000s

... and is thus the **fastest technology to be adapted** in **human history**.

Source: Flurry Research

Wearable Devices



iWatch



Samsung Gear



I'm Watch



Fitbit

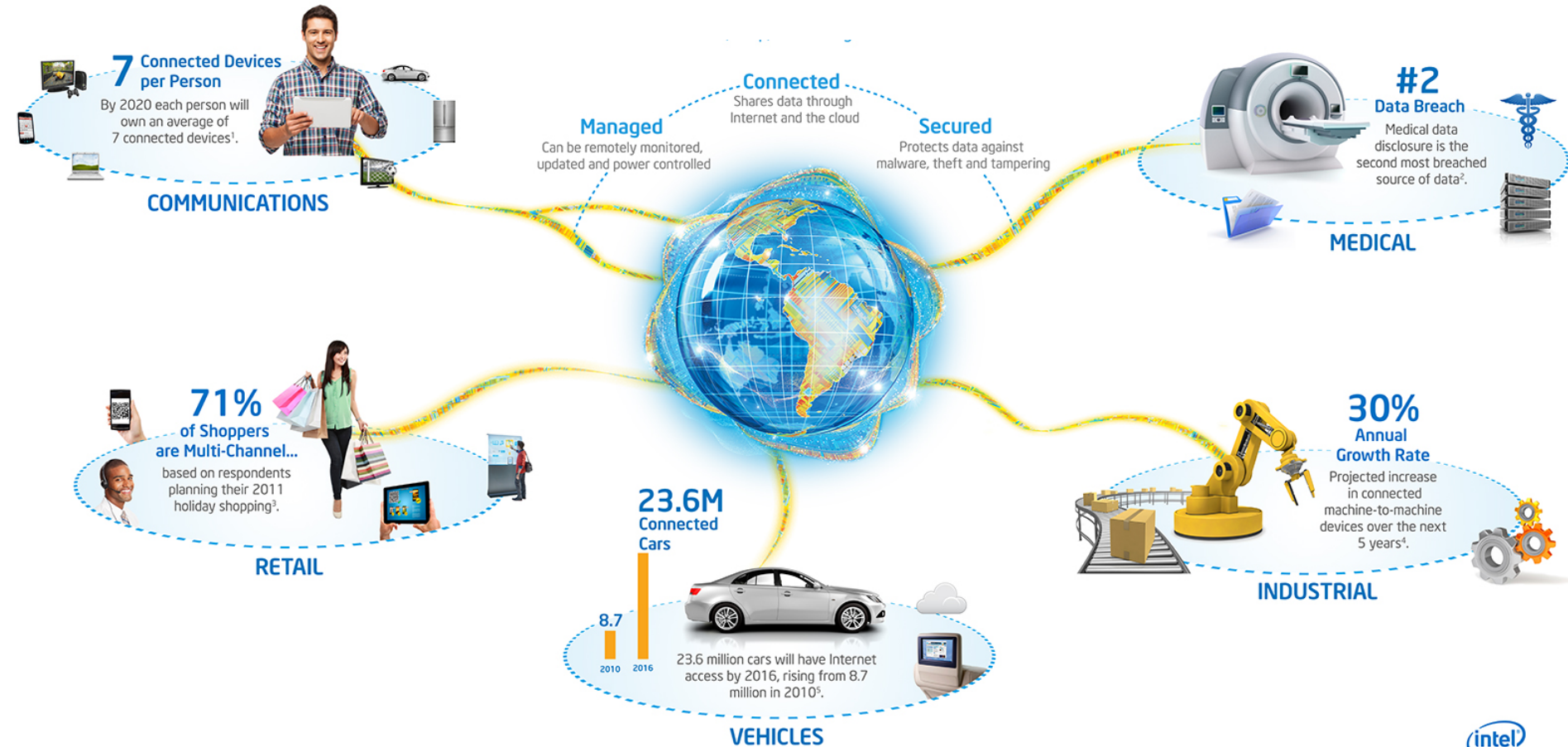


Fuelband



Google Glass

Internet of Things



Big Data Explosion

Digitalization creates massive amounts of data (Gartner 2010)

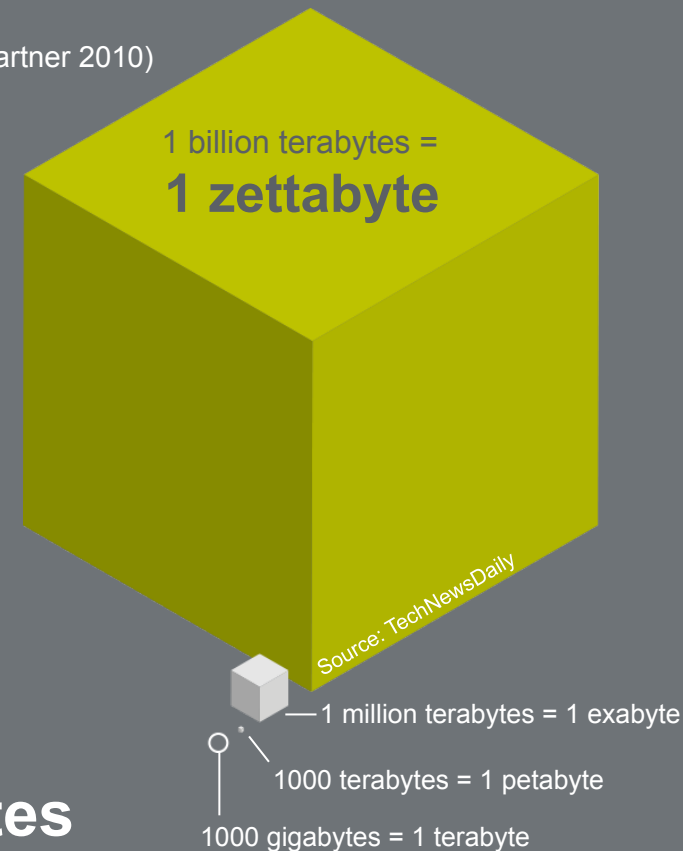
In 2016 worldwide: 4.1 Zettabytes (Sextillion: 21 zeros)

How long does it mankind take to create 5 Exabyte of data

- until 2000: **2000 Years**
- in 2011: **Two Days**
- in 2013: **Ten Minutes**

Source: C. Humbly

4.100.000.000.000.000.000.000 Bytes



Are we able to handle this amount of data? YES!

Example:

1 billion documents with 4 KB text each
(~ 1 standard page of text) = 4 TB memory

Read 4 TB of data from the hard drive:

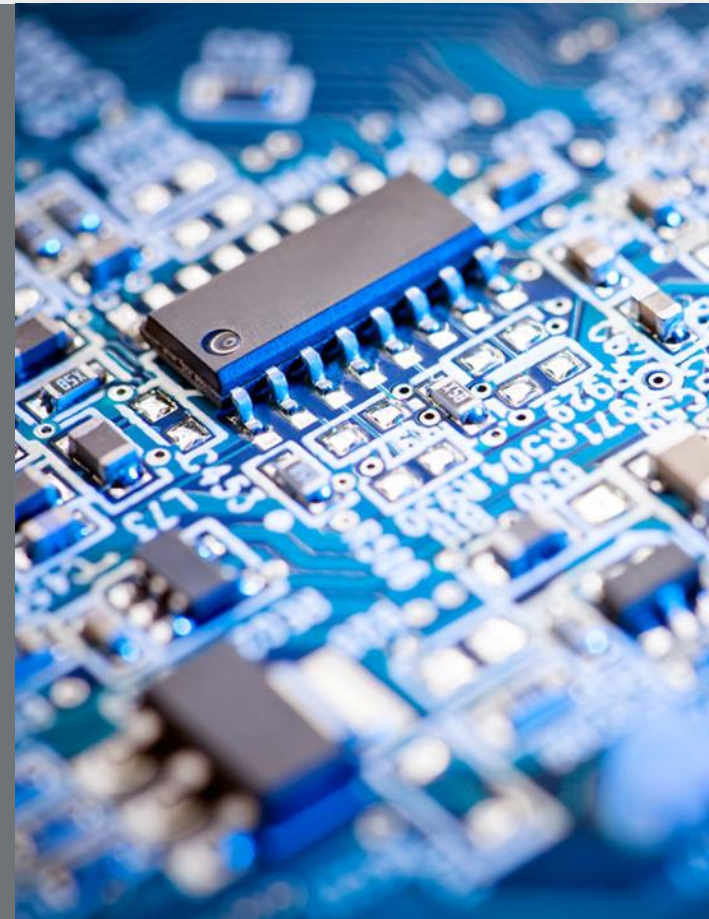
~ **1 ½ Day**

Read 4 TB from RAM:

~ **15 minutes**

Costs for 4 TB RAM

- **2002: \$ 600,000**
- **2007: \$ 100,000**
- **2012: \$ 20,000**



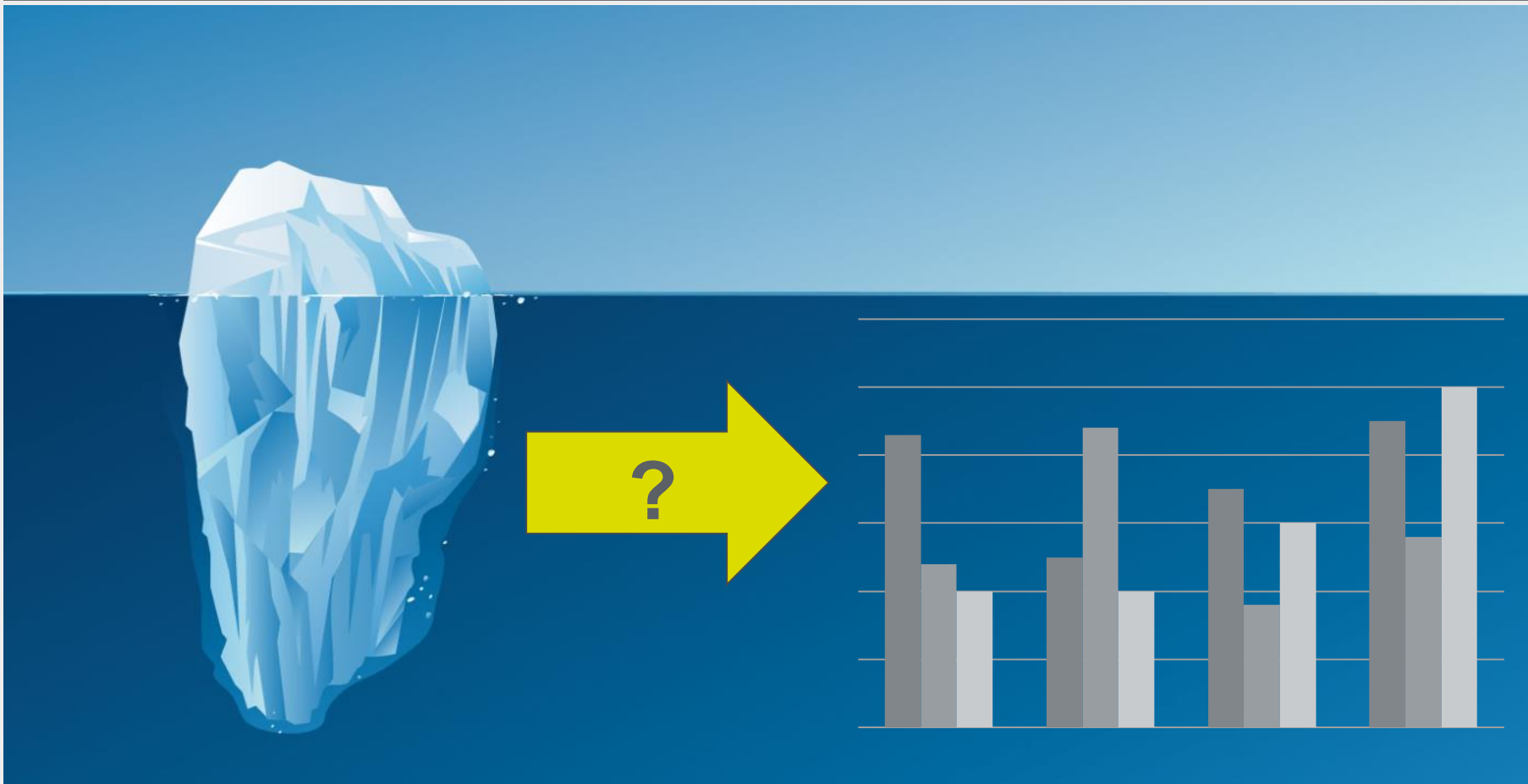
BUT: Only a small amount of the available information is easily and immediately accessible!



Only 20% of relevant information is structured, e.g. in databases or Excel spreadsheets, and is readily available

80% of critical information is unstructured or only minimally structured, e.g. documents, emails, Power Point slides, PDFs, and often difficult to access and not readily available

Challenge: Insights into Unstructured Data



McKinsey&Company

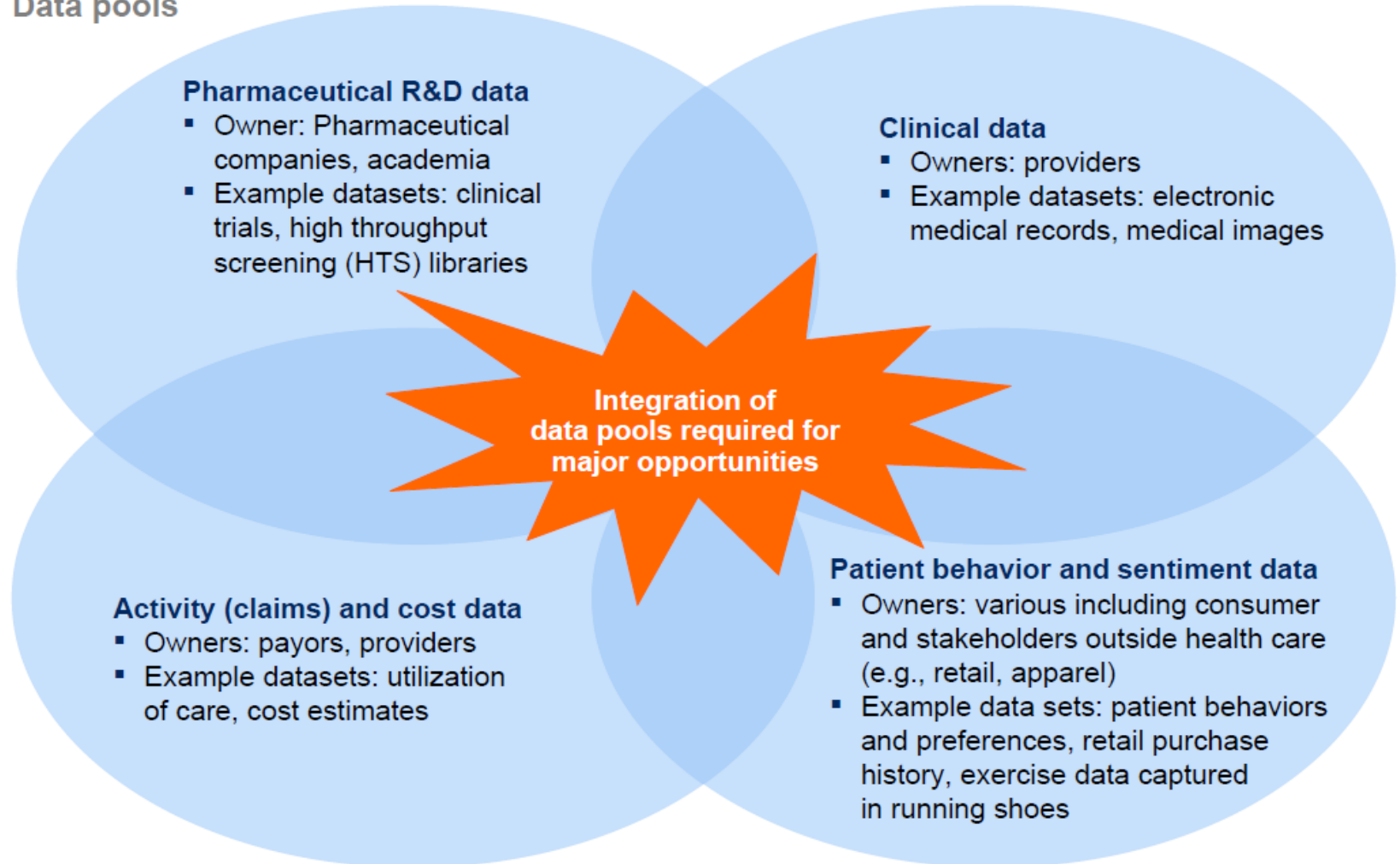
McKinsey Global Institute



June 2011

Big data: The next frontier
for innovation, competition,
and productivity

Data pools



SOURCE: McKinsey Global Institute analysis

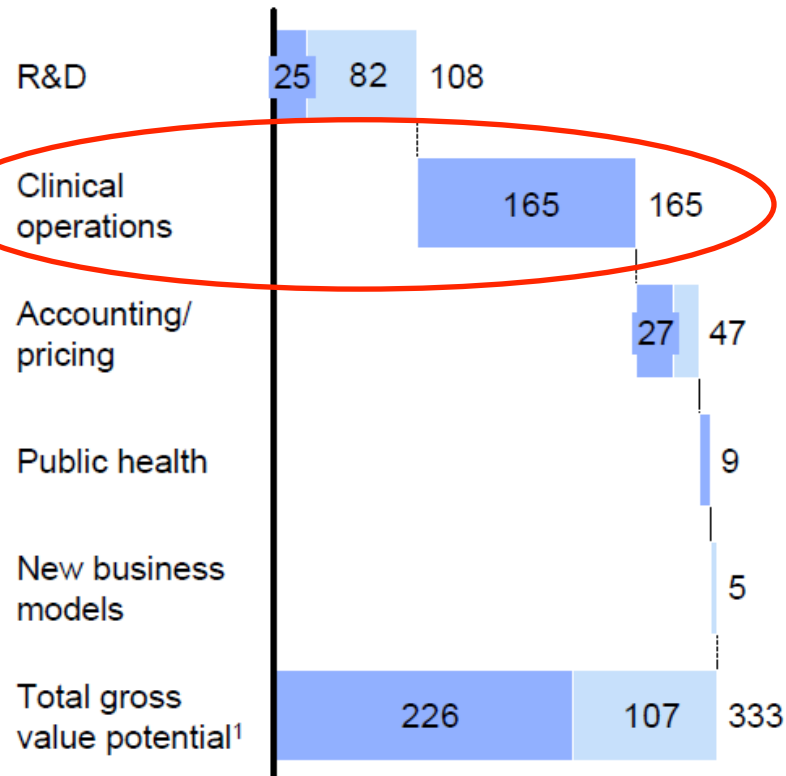
The estimated long-term value of identified levers is more than \$300 billion, with potentially more than \$200 billion savings on national health care spending

Value potential from use of big data

\$ billion per year

- Direct reduction on national health care expenditure
- Unclear impact on national health care expenditure

Lever examples



Predictive modeling to determine allocation of R&D resources, clinical trial design, and personalized medicine

Comparative effectiveness research (CER), clinical decision support system, and dashboards for transparency into clinical data

Advanced algorithms for fraud detection, performance-based drug pricing

Public health surveillance and response systems

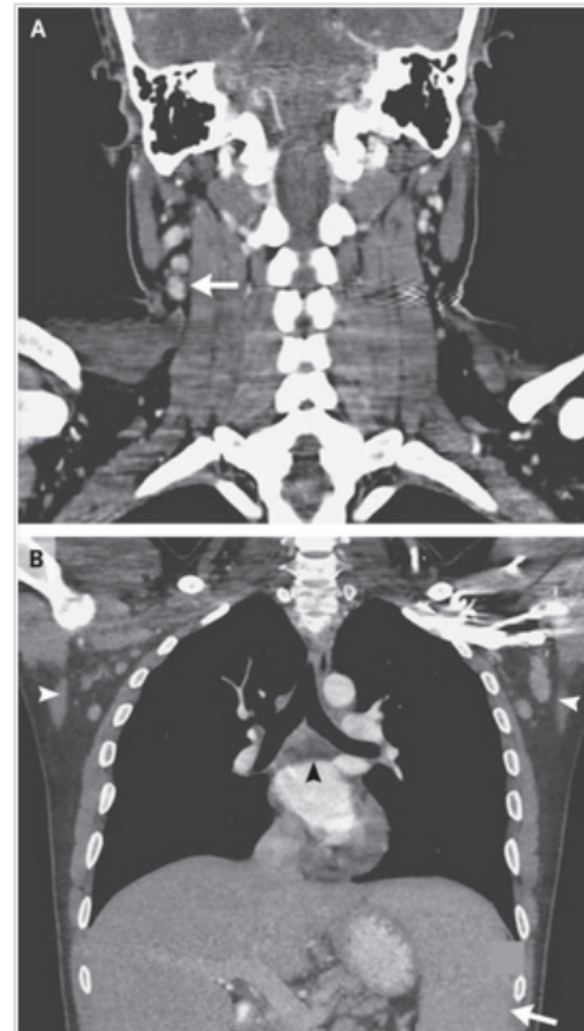
Aggregation of patient records to provide datasets and insights; online platforms and communities

Example: Analysis of Patient Records

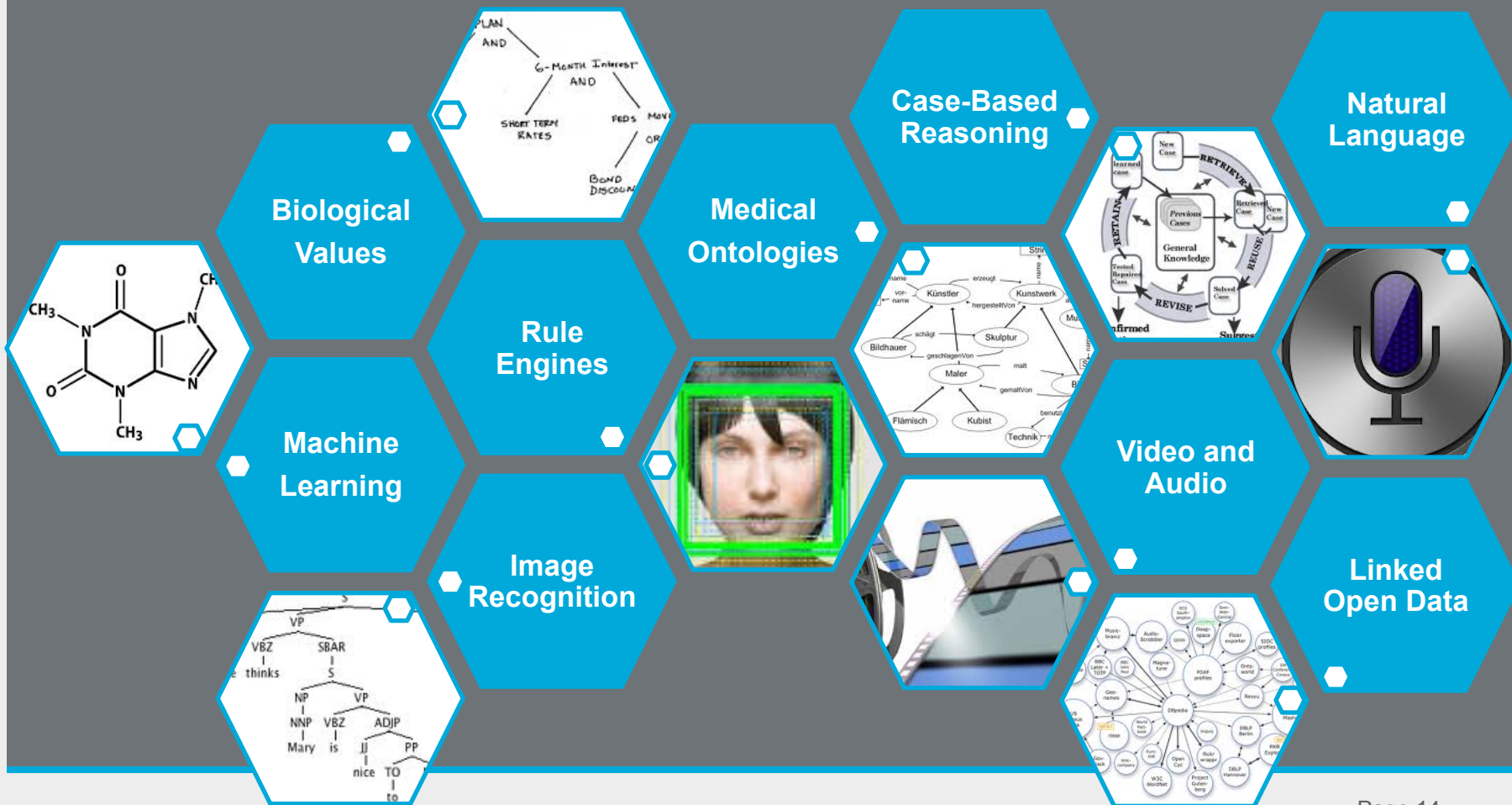
and a dull headache and was admitted to the hospital. He reported no night sweats, weight loss, rhinorrhea, neck rigidity, neck pain, vision changes, weakness, or numbness and no respiratory, gastrointestinal, or genitourinary symptoms.

On examination, the blood pressure was 106/58 mm Hg, and the pulse 104 beats per minute; the respirations were 20 breaths per minute, with an oxygen saturation of 98% while the patient was breathing ambient air. There was a superficial ulcer on the right side of the lower lip, and the spleen tip was palpable approximately 3 cm below the costal margin; the remainder of the examination was unchanged. Blood levels of lactic acid, C3, C4, haptoglobin, creatine kinase, and thyrotropin were normal; testing for rheumatoid factor and antinuclear antibodies was negative. Other test results are shown in [Table 1](#). Urinalysis revealed yellow, clear fluid, with 1+ occult blood, trace albumin, and few squamous cells per high-power field, and was otherwise normal. Fluids were administered intravenously, and another blood sample was obtained for culture. The maximal daily temperature was 40.7°C on the first day. On the second day, the temperature rose to 40.8°C.

Figure 2. CT Scans of the Neck and Chest.



To process Medical Data a Broad Scope of Technologies is needed



Linguistics and Natural Language Processing

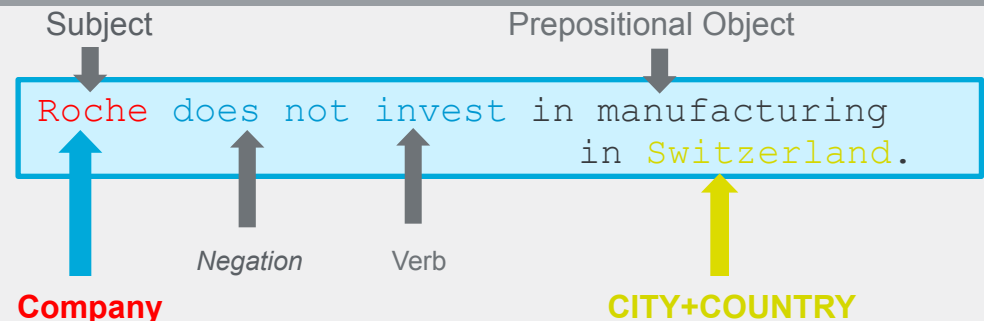


Linguistic and semantic technologies help better grasp the “meaning” of information

Cutting-edge, patented linguistics and semantics

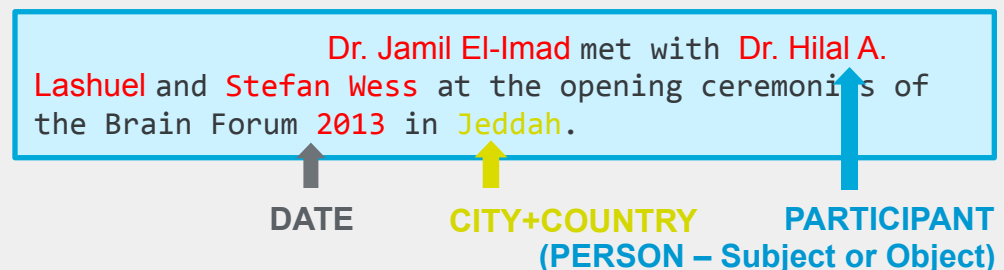
- Stemming
- Compound words
- Deep morphosyntactical analysis
- Sentiment analysis
- Person aliasing: replacing pronouns with the correct name
- Entities
 - >20 out-of-the-box
 - ontologies
- Facts
- 32 languages
- Configurable, adjustable and expandable via SALSA and ontologies

Deep morphosyntactical analysis, out-of-the-box entities

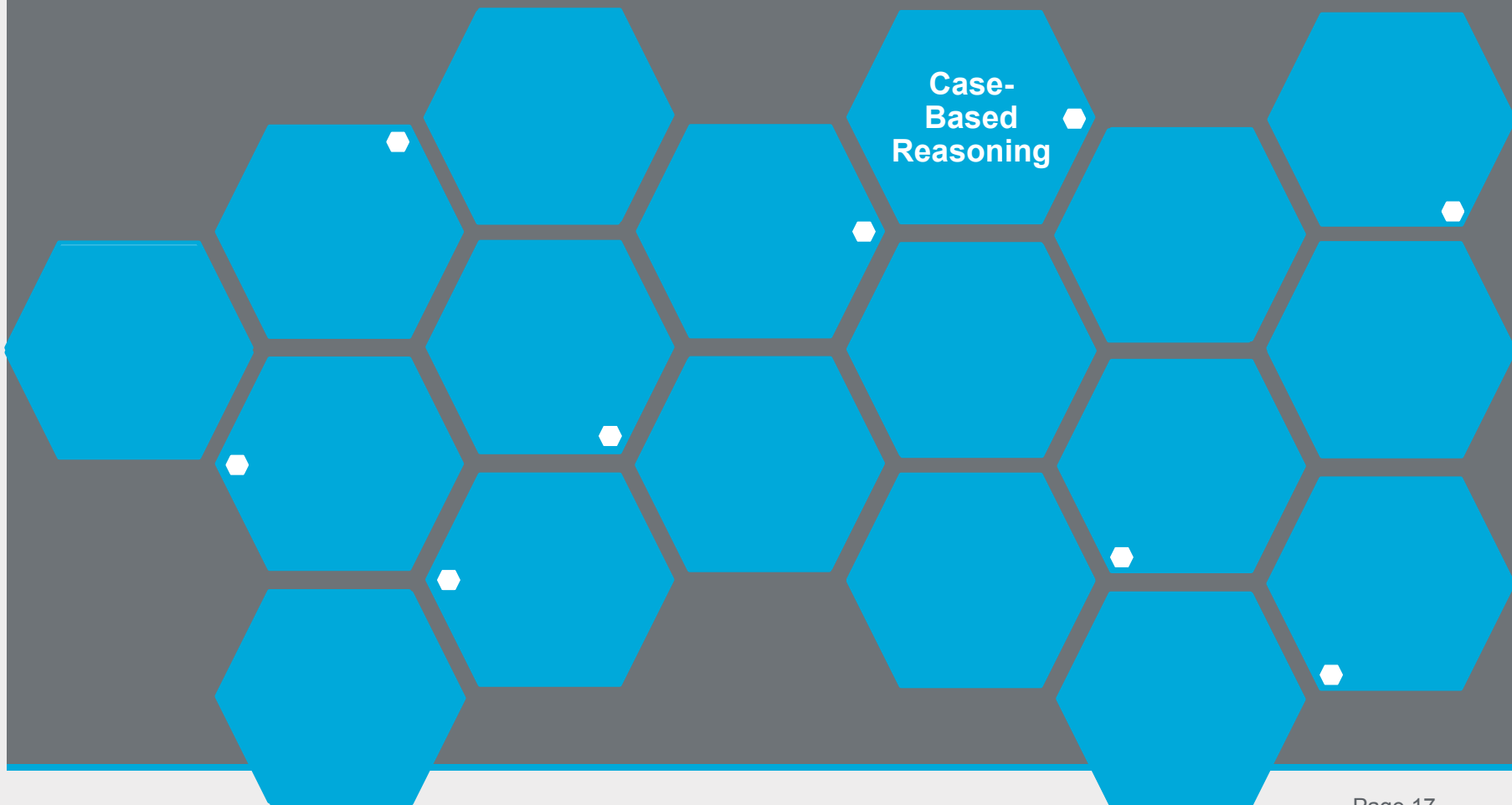


Application-specific fact (example)::

Meeting := <Participants, CITY, DATE>



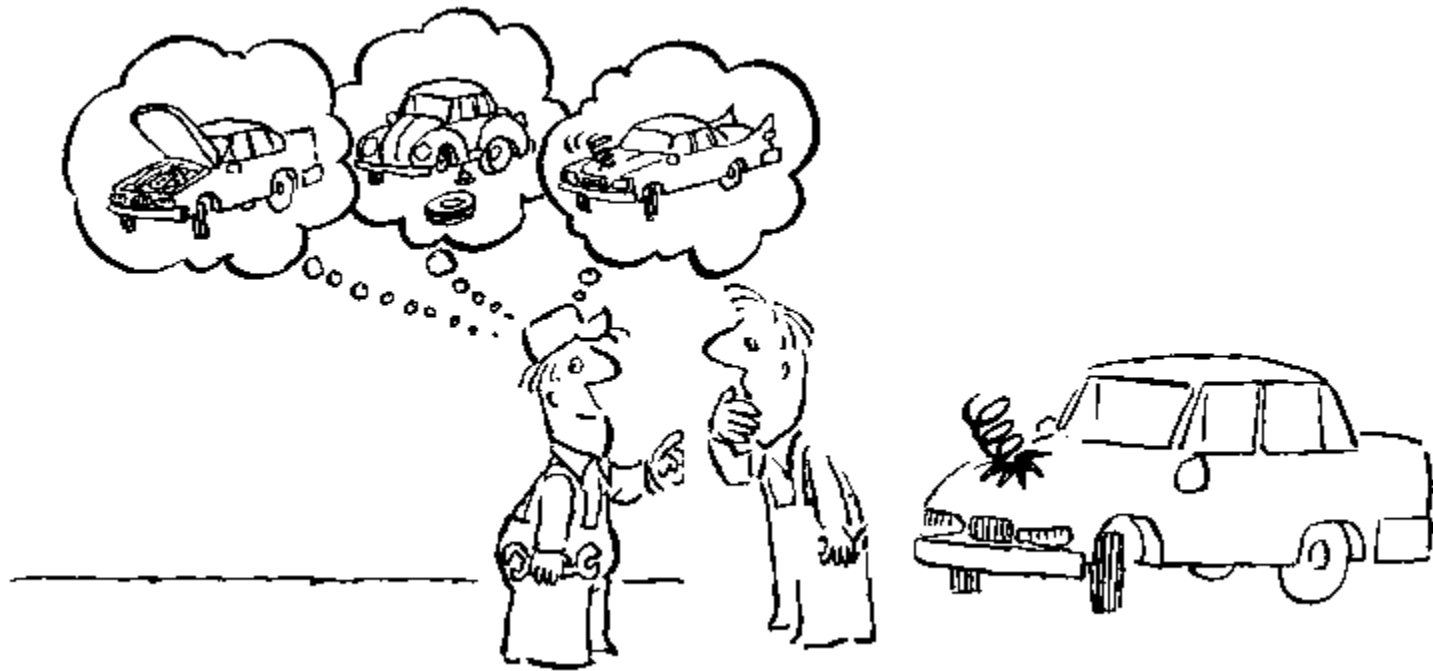
Case-Based Reasoning



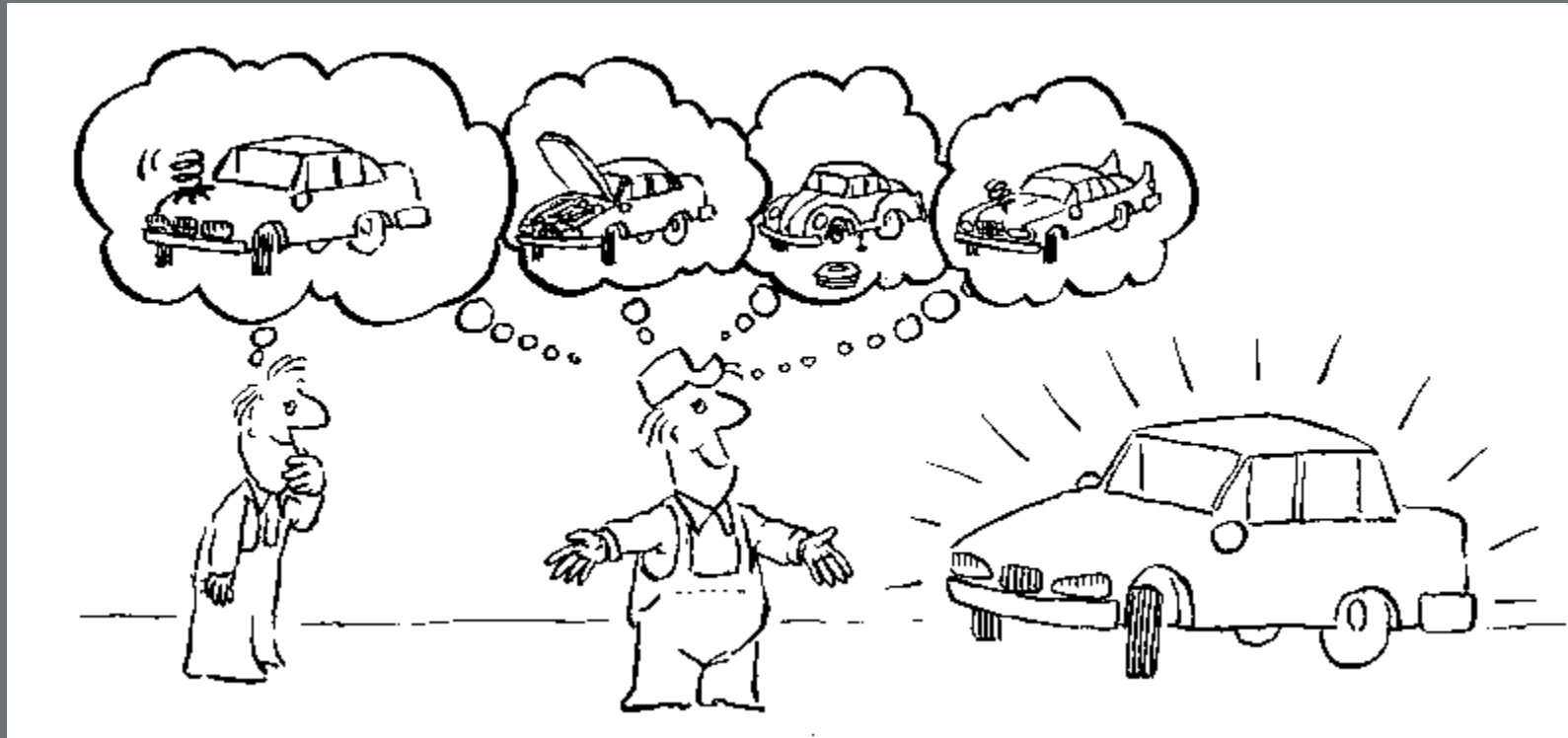
Case Based Reasoning (CBR) – Idea: A new problem arises...



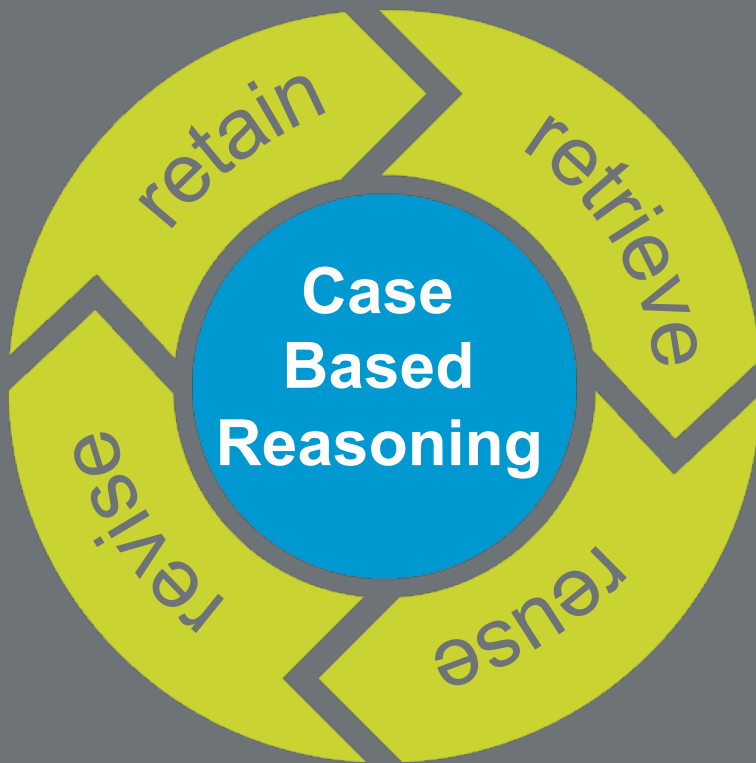
Case Based Reasoning (CBR) – Idea: ... has there been a similar problem before?...



Case Based Reasoning (CBR) – Idea: ... Solution found, Experience growing!

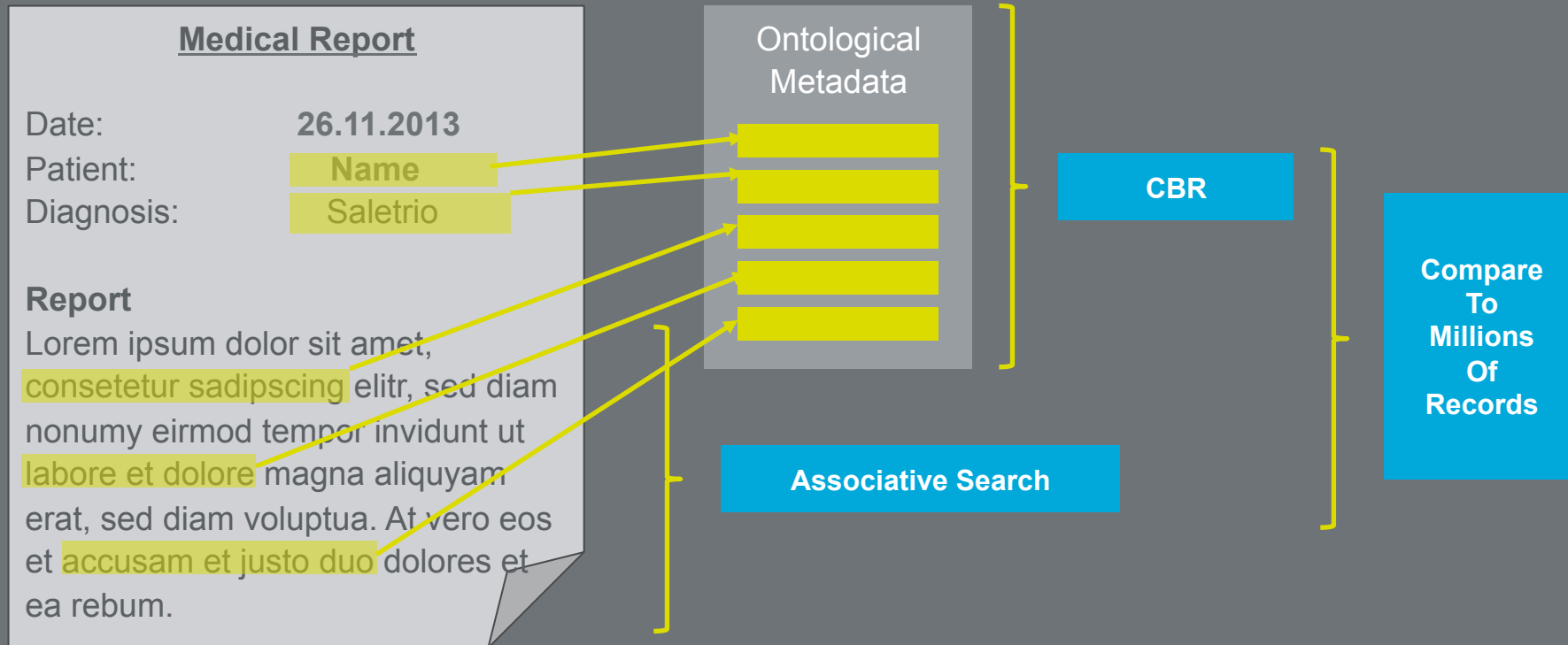


Case Based Reasoning: Solving new problems by re-using known cases

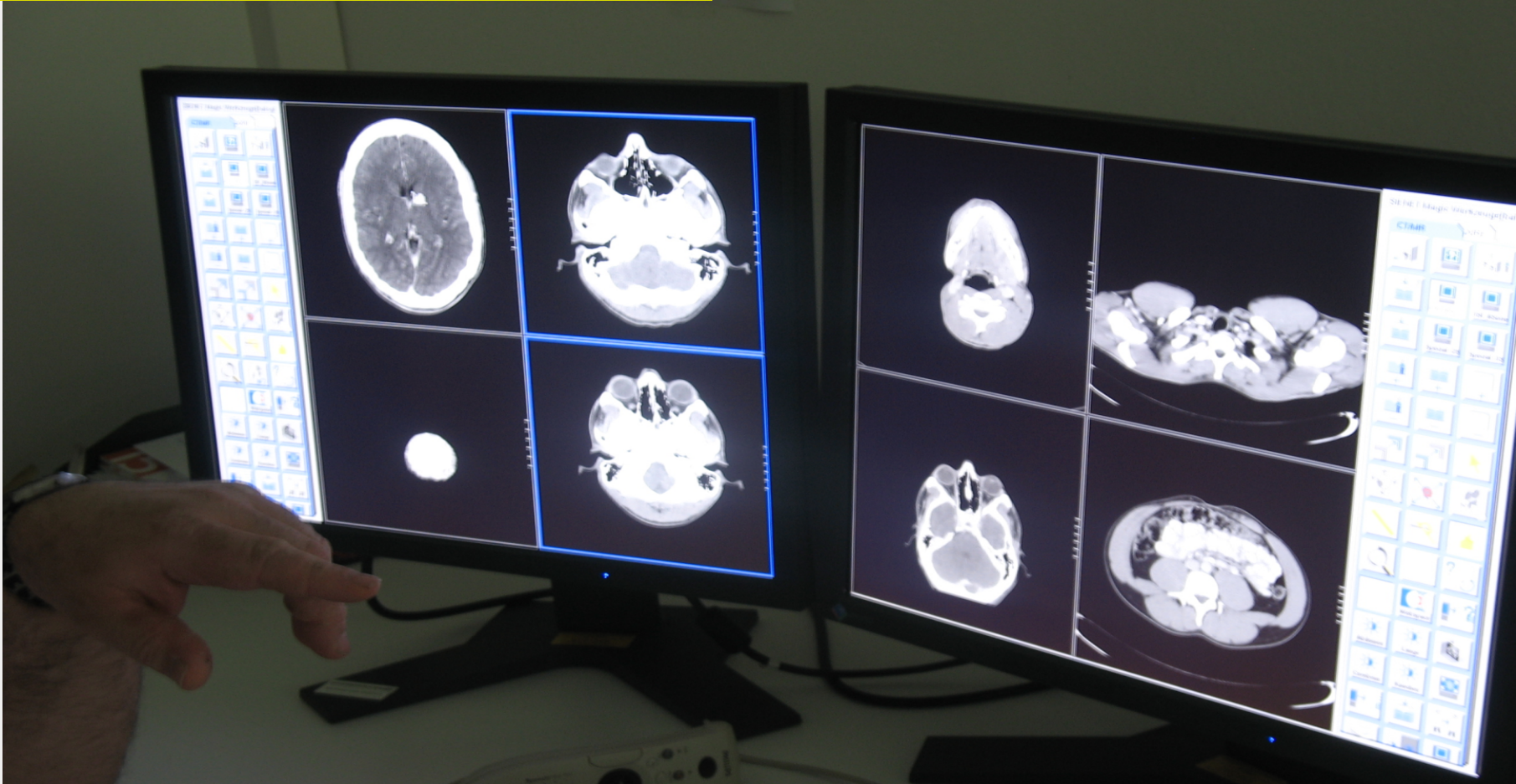


CBR traces its roots to the work of Roger Schank and his students at Yale University in the early 1980s. Schank's model of dynamic memory was the basis for the earliest CBR system: Janet Kolodner's CYRUS in 1985.

Both Paradigms Are Seamlessly Integrated



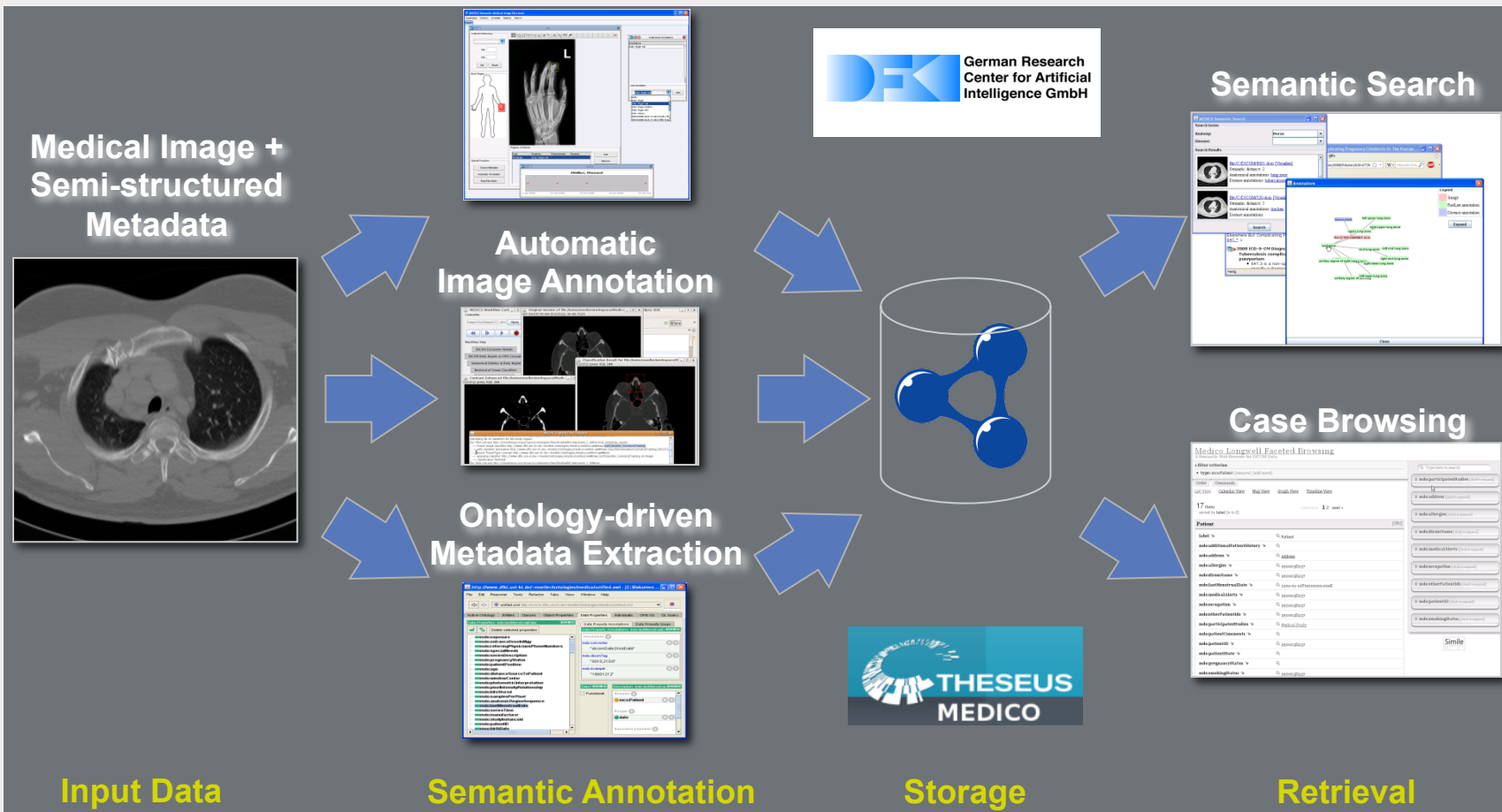
THESEUS Medico



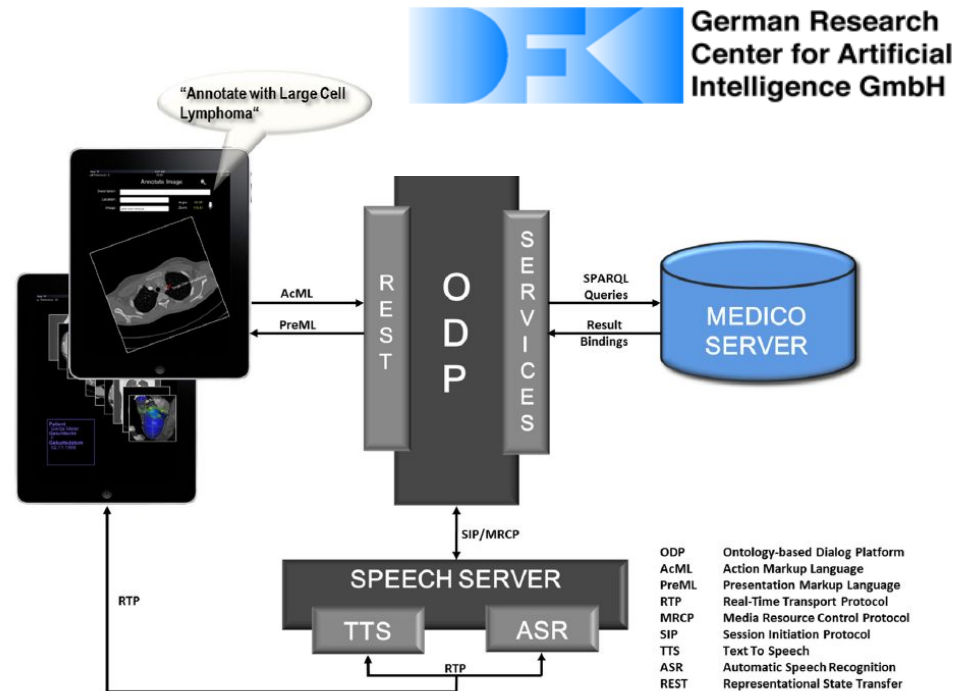
SIEMENS



Medico implements a semantic analysis workflow of medical images



RadSpeech



German High Tech Award, 2011

Trend : Monitor own health condition 24/7



The ongoing consumerization of health-care will allow an ongoing monitoring of the personal health condition. The “quantified self” movement will become main stream.

Self monitoring will change negative health related behaviors. Technology will allow to predict health related problems very early. The collected health data will help to speed up medical research.

Trend : Get the best, personalized treatment



All medical literature and research is available online. Digitization of medical records and ongoing self monitoring will create a unique and large experience base of health related issues and treatments.

Smart technologies will allow to search and compare all existing medical cases worldwide helping doctors finding best treatments in real time.