



GS05 – Become an Analytics and Cognitive Business with IBM z Systems

Leipzig, 24. October 2016

Dr. Manfred Gnirss
IBM Client Center Boeblingen
Germany

Wilhelm Mild IBM R&D Lab. Germany

10th European GSE/IBM Technical University for z/VSE, z/VM, KVM and Linux on IBM z Systems



October 24th – 26th, 2016 The Westin Hotel in Leipzig, Germany





Trademarks

The following are trademarks of the International Business Machines Corporation in the United States and/or other countries.

DB2*	ECKD	IBM*	LinuxONE	PR/SM	z13	z Systems
DB2 Connect	FICON*	Ibm.com	LinuxONE Emperor	Storwize*	zEnterprise*	z/VSE*
DS8000*	FlashSystem	IBM (logo)*	LinuxONE Rockhopper	XIV*	z/OS*	z/VM*

^{*} Registered trademarks of IBM Corporation

The following are trademarks or registered trademarks of other companies.

Adobe, the Adobe logo, PostScript, and the PostScript logo are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States, and/or other countries.

Cell Broadband Engine is a trademark of Sony Computer Entertainment, Inc. in the United States, other countries, or both and is used under license therefrom.

Intel, Intel logo, Intel Inside, Intel Inside logo, Intel Centrino, Intel Centrino logo, Celeron, Intel Xeon, Intel SpeedStep, Itanium, and Pentium are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

IT Infrastructure Library is a registered trademark of the Central Computer and Telecommunications Agency which is now part of the Office of Government Commerce.

ITIL is a registered trademark, and a registered community trademark of the Office of Government Commerce, and is registered in the U.S. Patent and Trademark Office.

Java and all Java based trademarks and logos are trademarks or registered trademarks of Oracle and/or its affiliates.

Linear Tape-Open, LTO, the LTO Logo, Ultrium, and the Ultrium logo are trademarks of HP, IBM Corp. and Quantum in the U.S. and

Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.

Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

OpenStack is a trademark of OpenStack LLC. The OpenStack trademark policy is available on the OpenStack website.

TEALEAF is a registered trademark of Tealeaf, an IBM Company.

Windows Server and the Windows logo are trademarks of the Microsoft group of countries.

Worklight is a trademark or registered trademark of Worklight, an IBM Company.

UNIX is a registered trademark of The Open Group in the United States and other countries.

Notes:

Performance is in Internal Throughput Rate (ITR) ratio based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve throughput improvements equivalent to the performance ratios stated here.

IBM hardware products are manufactured from new parts, or new and serviceable used parts. Regardless, our warranty terms apply.

All customer examples cited or described in this presentation are presented as illustrations of the manner in which some customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics will vary depending on individual customer configurations and conditions.

This publication was produced in the United States. IBM may not offer the products, services or features discussed in this document in other countries, and the information may be subject to change without notice. Consult your local IBM business contact for information on the product or services available in your area.

All statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.

Information about non-IBM products is obtained from the manufacturers of those products or their published announcements. IBM has not tested those products and cannot confirm the performance, compatibility, or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

Prices subject to change without notice. Contact your IBM representative or Business Partner for the most current pricing in your geography.

This information provides only general descriptions of the types and portions of workloads that are eligible for execution on Specialty Engines (e.g., zIIPs, zAAPs, and IFLs) ("SEs"). IBM authorizes customers to use IBM SE only to execute the processing of Eligible Workloads of specific Programs expressly authorized by IBM as specified in the "Authorized Use Table for IBM Machines" provided at

www.ibm.com/systems/support/machine_warranties/machine_code/aut.html ("AUT"). No other workload processing is authorized for execution on an SE. IBM offers SE at a lower price than General Processors/Central Processors because customers are authorized to use SEs only to process certain types and/or amounts of workloads as specified by IBM in the AUT.

^{*} Other product and service names might be trademarks of IBM or other companies.





Cognitive Computing

Analytics

= Digital Business

Cognitive Computing = Digital Business + Digital Intelligence





World's leading businesses run on the mainframe



92 of the top 100 worldwide banks



10
out of 10 of the world's largest insurers



23 of the top 25 US retailers



out of 25 of the world's largest airlines

Processing the world's transactions & data

30 billion

business transactions processed on the mainframe per day

91 percent

of surveyed CIOs said that new customer-facing applications are accessing the mainframe

80 percent

of the world's corporate data resides or originates on mainframes

55 percent

of all enterprise applications need the mainframe to complete transactions





The business environment is shifting...



...and a new era of computing is emerging





The business environment is shifting...



...and a new era of computing is emerging

Cognitive Computing = Digital Business + Digital Intelligence



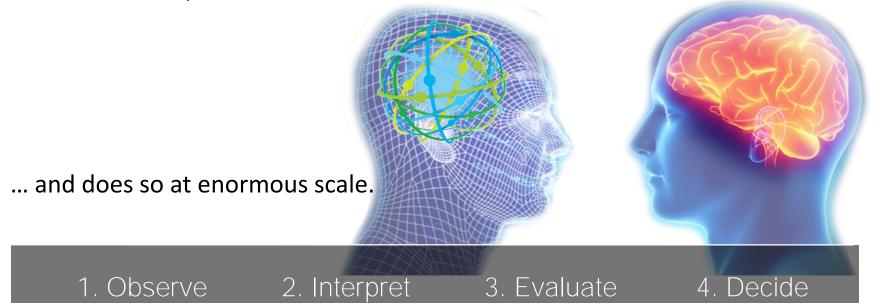
https://www.youtube.com/watch?v=AGYkk5wUKtA





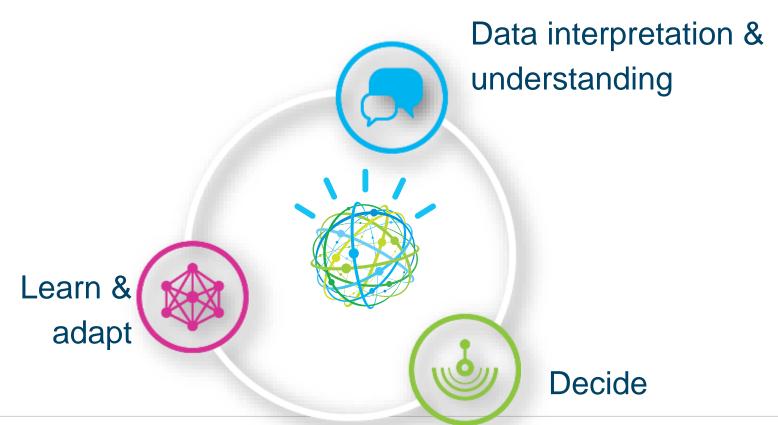
What is Cognitive Computing?

- Cognitive systems are able to learn their behavior through education
- That supports forms of expression that are more natural for human interaction
- Whose primary value is their expertise; and
- That continue to evolve as they experience new information, new scenarios, and new responses



What is Cognitive Computing?

Cognitive: Psychological processes involved in acquisition and understanding of knowledge, formation of beliefs and attitudes, and decision making and problem solving.

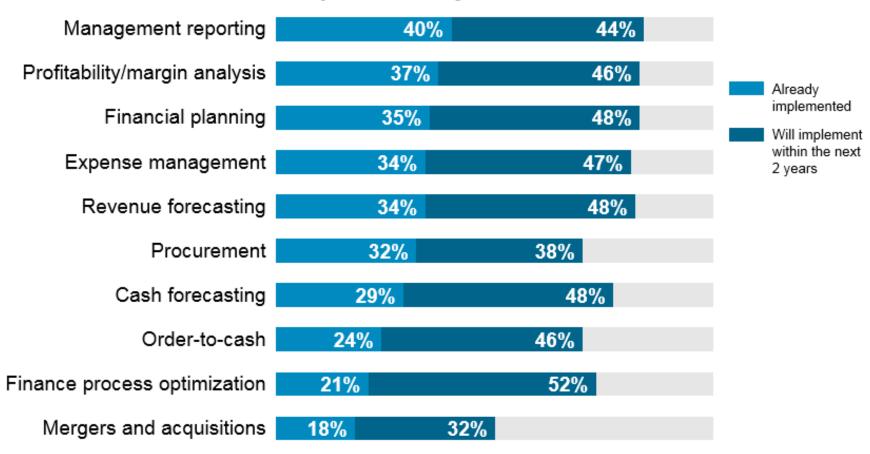




IBM Center for Applied Insights

http://ibmcai.com/

Adoption of analytics



Source: IBM Center for Applied Insights survey of 337 Finance executives across industries, July 2015.



From transactional data to real-time Analytics and Cognitive consider end-to-end solutions and operational impacts

Real-time "integration of analytics and transaction processing" increases customer value with every interaction

- Deliver real-time insights at the point of impact
- Manage data lifecycle and governance
- Eliminate redundancy and avoid ETL



IBM Software examples

- Cognos BI
- SPSS
- SPARK
- DB2
- InfoSphere® Warehouse
- DB2 Analytics Accelerator
- InfoSphere Information Server
- InfoSphere Data Replication
- InfoSphere Master Data Mgmt
- DB2
- IMS, VSAM
- Non IBM, e.g. Oracle

"Cognos generates insightful reports and sophisticated dashboards, providing quick and accurate information to senior management. We are now adding more reporting functionality - on business revenue, credit data, loan risks, and so on - to make Cognos the complete decision-support system for Sicoob."

- Paulo Nassar,

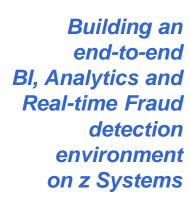
IT Processing and Storage Infrastructure Manager, Sicoob

IBM Cognos Business Intelligence and additional analytics software is running on Linux on z Systems



Data: From Database to Information Management







SPARC Cognos BI **SPSS InfoSphere** Warehouse **DB2 LUW / Linux DB2 BLU** DB2 z/OS

DB₂ Warehouse **Database**

IBM TotalStorage / Flash

- **IBM DB2 Analytics Accelerator**
- **Federation Server**

MS

SQL

VSAM ORACLE IMS data

Linux on z Systems

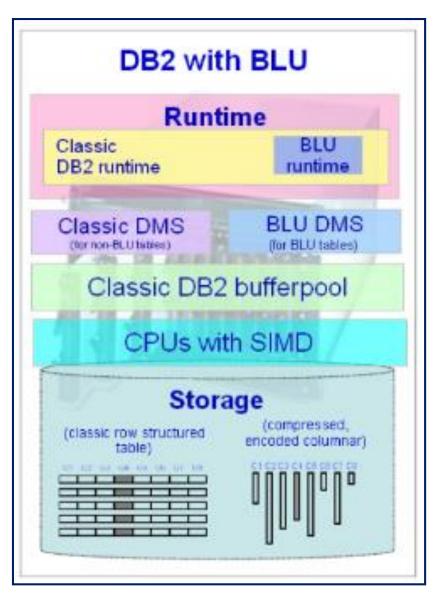
- **BI** solutions **SPARC**
- **DB2 BLU**

Accelerator on z Systems

BLU Acceleration for Linux on z Systems

Super Simple. Super Fast.





Solution

- DB2 with BLU Acceleration is the preferred solution for customers who would like to run analytics on z Systems Linux data
- Satisfy requirement for a columnar in-memory db
- Alternative of Linux on z
 Oracle installations
- Enhanced for distributed consolidations onto z Systems

Analytical:

From Transactional Workloads to Analytic Workloads

Two VERY different requirements for storing and processing data

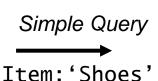
Customer



Business Transaction



Transactional Database



Cost: '\$34'

Cust: 'James'

2	2015 Sales	
	Cost	Cust

ltem	Cost	Cust
Shoes	\$34	James

Business Analyst



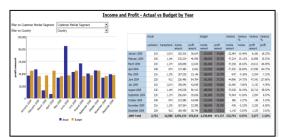
Complex Query

Sales & Profit for Shoes & Belts Year >= 2010

Data Warehouse



BI Reports & Dashboards





Business Analytics Solutions on z Systems

Business analytics capabilities



Cognos – Business Intelligence





SPSS – Predictive Analytics



Collaboration and Deployment Services

Business outcomes/benefits

- Understand current & potential state
- Monitor results & fine-tune your business
- Inform strategy with a view into the future
- Predict customer segment & category affinity
- Market Basket Analysis to identify NBO
- Overlay browsing history onto purchase history to profile customers



TM1 – Performance Management



Reporting, analysis, operational & financial planning and consolidation

- Product profitability across customers, business& channels
- Sales Performance Management to improve efficiency in incentive compensation process



Hadoop - Investigative Analytics



- Gain additional insights from LOGs, social media, streams, machine data, mass archives
- Understand and visualize the context of data in unstructured documents, LOGs and understand customer sentiment using Hadoop

Deutsche Region Clude NoSQL Data:



IBM Cloudant

IBM Cloudant & IBM's Big Data Portfolio

IBM is unique in having developed an enterprise class big data and analytics platform that allows you to address the full spectrum of big data business challenges. Cloudant provides another leading solution to the already market leading portfolio.

Cloudant - is a NoSQL database platform built for the cloud.

- on Linux on z Systems integration with IBM Big Data Products:

SPARK using Cloudant data for analytics

DB2 BLU **is IBM's in**-memory high performance relational database system (RDBMS) for analytics

- Cloudant is complementary to DB2 BLU
- ➤ Data from Cloudant can be loaded into DB2 BLU directly for analytics or via the BigInsights landing zone as part of Watson Foundations



Effective Analytics: Apache SPARK using Hadoop

http://hadoop.apache.org/



Search with Apache Solr

Search

Last Published: 10/31/2015 05:08:14

Welcome to Apache™ Hadoop®!



What Is Apache Hadoop?

The Apache™ Hadoop® project develops open-source software for reliable, scalable, distributed computing.

The Apache Hadoop software library is a framework that allows for the distributed processing of large data sets across clusters of computers using simple programming models. It is designed to scale up from single servers to thousands of machines, each offering local computation and storage. Rather than rely on hardware to deliver high-availability, the library itself is designed to detect and handle failures at the application layer, so delivering a highly-available service on top of a cluster of computers, each of which may be prone to failures.

The project includes these modules:

- Hadoop Common: The common utilities that support the other Hadoop modules.
- Hadoop Distributed File System (HDFS™): A distributed file system that provides high-throughput access to application data.
- Hadoop YARN: A framework for job scheduling and cluster resource management.
- Hadoop MapReduce: A YARN-based system for parallel processing of large data sets.

Other Hadoop-related projects at Apache include:

- Ambari™: A web-based tool for provisioning, managing, and monitoring Apache Hadoop clusters which includes support for Hadoop HDFS, Hadoop MapReduce, Hive, HCatalog, HBase, ZooKeeper, Oozie, Pig and Sgoop. Ambari also provides a dashboard for viewing cluster health such as heatmaps and ability to view MapReduce, Pig and Hive applications visually alongwith features to diagnose their performance characteristics in a user-friendly manner.
- Avro™: A data serialization system.
- Cassandra™: A scalable multi-master database with no single points of failure.
- Chukwa™: A data collection system for managing large distributed systems.
- HBase™: A scalable, distributed database that supports structured data storage for large tables.
- HiveTM: A data warehouse infrastructure that provides data summarization and ad hoc querying.
- Mahout™: A Scalable machine learning and data mining library.
- Piq™: A high-level data-flow language and execution framework for parallel computation.
- Spark™: A fast and general compute engine for Hadoop data. Spark provides a simple and expressive programming model that supports a wide range of applications, including ETL, machine learning, stream processing, and graph computation.
- <u>Tez</u>^m: A generalized data-flow programming framework, built on Hadoop YARN, which provides a powerful and flexible engine to execute an arbitrary DAG of tasks to process data for both batch and interactive use-cases. Tez is being adopted by Hive™, Pig™ and other frameworks in the Hadoop ecosystem, and also by other commercial software (e.g. ETL tools), to replace Hadoop™ MapReduce as the underlying execution engine.
- ZooKeeper™: A high-performance coordination service for distributed applications.



What Spark Is, What it Is Not

- An Apache Foundation open source project
- Not a product
- An in-memory compute engine that works with data
- Not a data store
- Enables highly sophisticated analysis on huge volumes of data at scale
- Unified environment for data scientists, developers and data engineers
- Radically simplifies the process of developing intelligent apps fueled by data



What Spark on z/OS is Not

What it is not

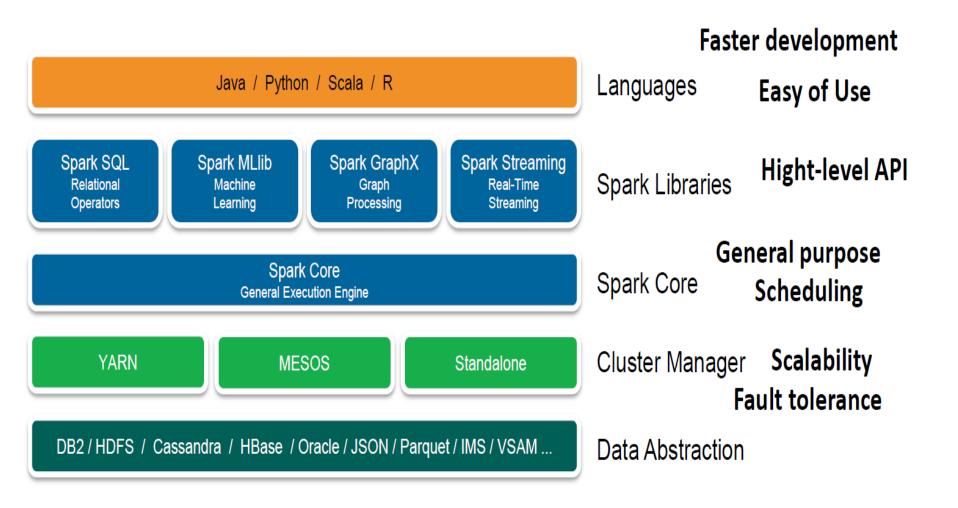
- •A data cache for all data in DB2, IMS, IDAA, VSAM ...
- Just a different SQL engine or query optimizer
- •An effective mechanism to access a *single* data source for analytics

Why isn't it the same as a query acceleration / IBM DB2 Analytics Accelerator?

- Spark does not optimize SQL queries
- Spark is not a mechanism to store data, it rather provides interfaces to uniformly access data & to apply analytics using a unified interface
- DB2 Analytics Accelerator interaction with applications is via DB2;
 Spark interaction with applications is via Spark interfaces (Stream, MLlib, Graphx, SQL), driven through REST or Java
- Spark analytics can access data in DB2, DB2 Analytics Accelerator, VSAM, IMS, off platform, etc.



The Spark Stack – compute engine for analytics Architectural Overview



BM Commitment:

IBM major Commitment to Advance Analytics with Apache® Spark™

06/2015 Announcment

https://www-03.ibm.com/press/us/en/pressrelease/47107.wss

 At the core of this commitment, IBM plans to embed Spark into its industryleading <u>Analytics</u> and <u>Commerce</u> platforms, and to offer Spark as a service on <u>IBM</u> <u>Cloud</u>.

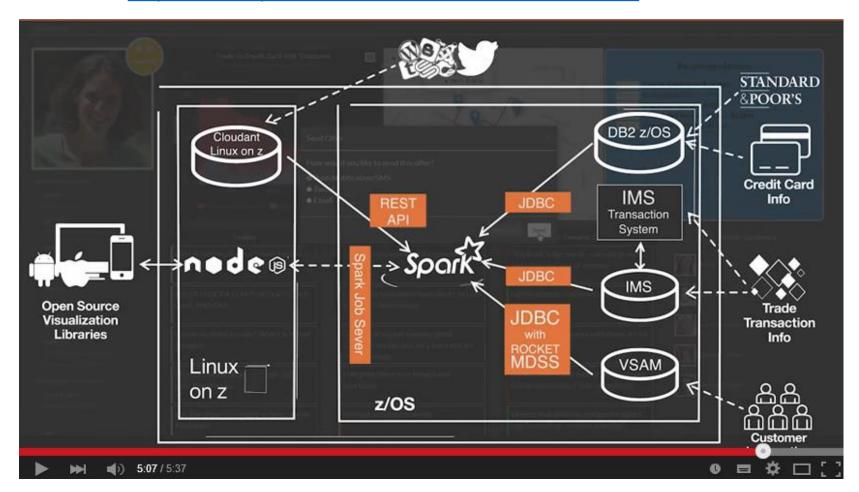
www.smartercomputingblog.com/system-z/ibm-spark-mainframe/

- Apache® Spark™ is an open-source computing framework with in-memory processing to speed analytic applications up to 100 times faster compared to technologies on the market today and enhance mission-critical applications with deep intelligence.
- The impact is being compared to that of Linux, Spark being a common framework for analytics just as Linux has been a common framework for computing.
- Linux and the mainframe have been a perfect match, and so it will be with
 Spark for both z/OS and <u>Linux on z Systems</u>.



Digital Business with SPARK: Spark Analytics on z Systems – the Technology Demo

https://www.youtube.com/watch?v=sDmWcuO5Rk8



Real time Analytics in combination with sentiment cloud data

- Performance
- with adatapian aspentruoriginal Macation BM z Systems

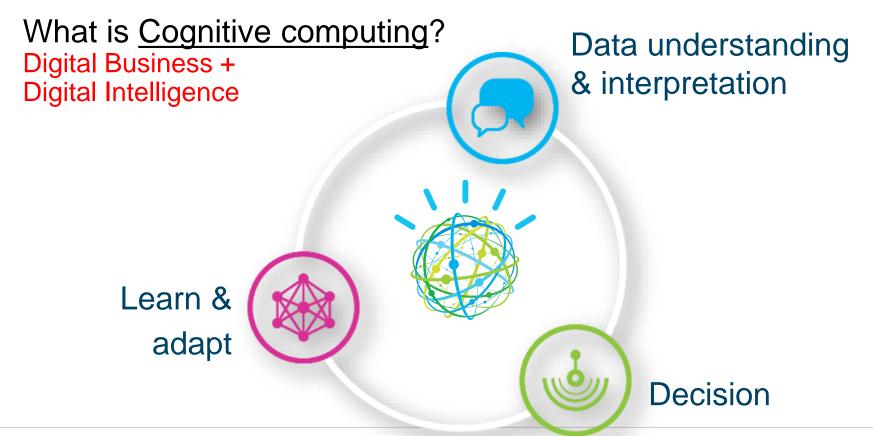


Watson technology – the Digital Intelligence for Cognitive Computing

- ➤ Watson is the first commercially available cognitive computing capability, representing a new era in computing.
- Watson analyzes high volumes of data and processes information more like a human than a computer
 - by understanding natural language,
 - generating hypotheses based on evidence,
 - > learning as it goes
- > IBM has created two business units:
 - Watson Group, established for the development and commercialization of cloud-delivered cognitive computing technologies
 - Watson Health to improve the ability of doctors, researchers and insurers
 - ➤ to surface new insights from the massive amount of personal health data being created daily to deliver personalized

What is <u>Cognitive</u>?

Cognitive: Psychological processes involved in acquisition and understanding of knowledge, formation of beliefs and attitudes, and decision making and problem solving.



Analytics = Digital Business

Real-time "integration of analytics and transaction processing" increases customer value with every interaction

- Deliver real-time insights at the point of impact
- Manage data lifecycle and governance
- Eliminate redundancy and avoid ETL



IBM Software examples

- Cognos BI
- SPSS
- SPARK
- DB2
- InfoSphere® Warehouse
- DB2 Analytics Accelerator
- InfoSphere Information Server
- InfoSphere Data Replication
- InfoSphere Master Data Mgmt
- DB2
- IMS, VSAM
- Non IBM, e.g. Oracle

"Cognos generates insightful reports and sophisticated dashboards, providing quick and accurate information to senior management. We are now adding more reporting functionality - on business revenue, credit data, loan risks, and so on - to make Cognos the complete decision-support system for Sicoob."

- Paulo Nassar,

IT Processing and Storage Infrastructure Manager, Sicoob

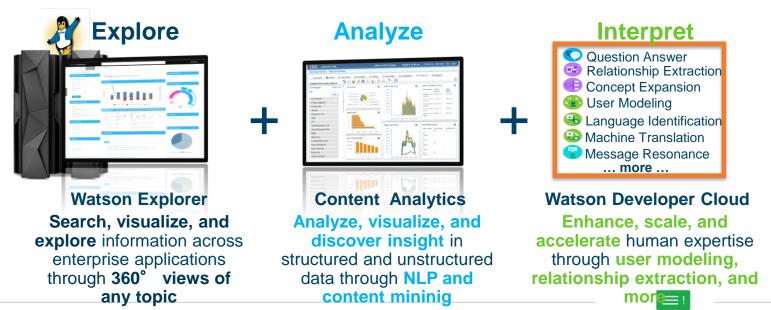
IBM Cognos Business Intelligence and additional analytics software is running on Linux on z Systems

Cognitive Computing = Digital Business + Digital Intelligence

The natural evolution beyond Analytics and digital business

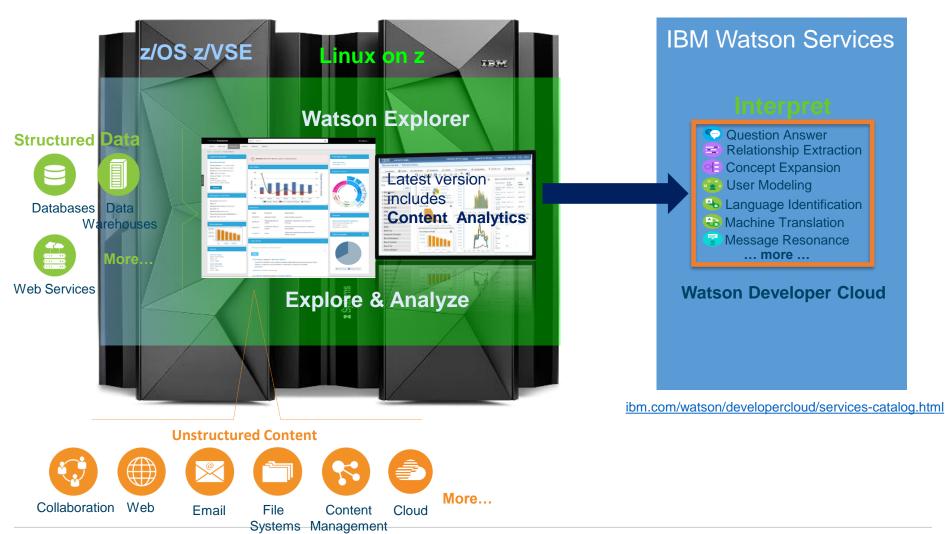
Watson technology – the Digital Intelligence for Cognitive Computing

- Watson is the first commercially available cognitive computing capability, representing a new era in computing.
- Watson analyzes high volumes of data and processes information more like a human than a computer
 - by understanding natural language,
 - generating hypotheses based on evidence,
 - learning as it goes
- ➤ The three pillars of Watson Explorer:



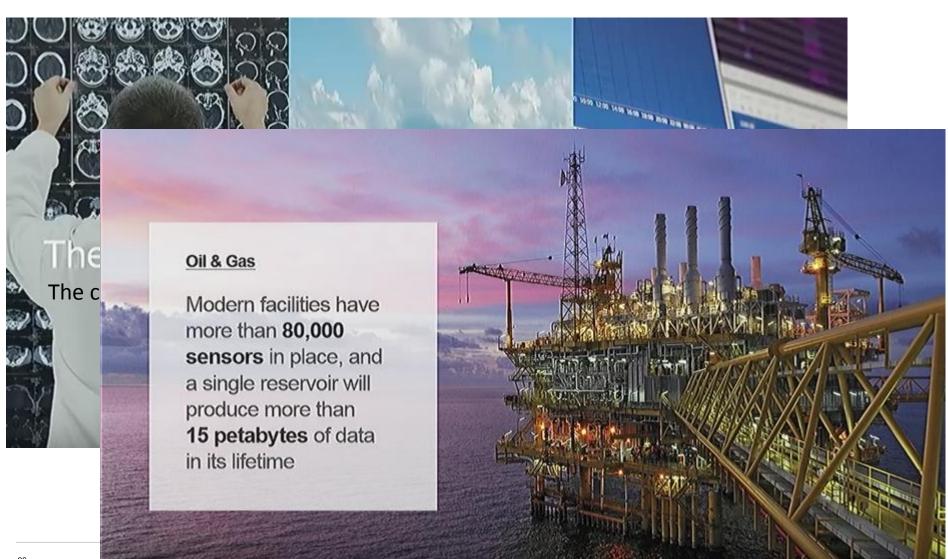
Watson Explorer is running on z Systems

Makes data from enterprise and non-enterprise silos easily accessible



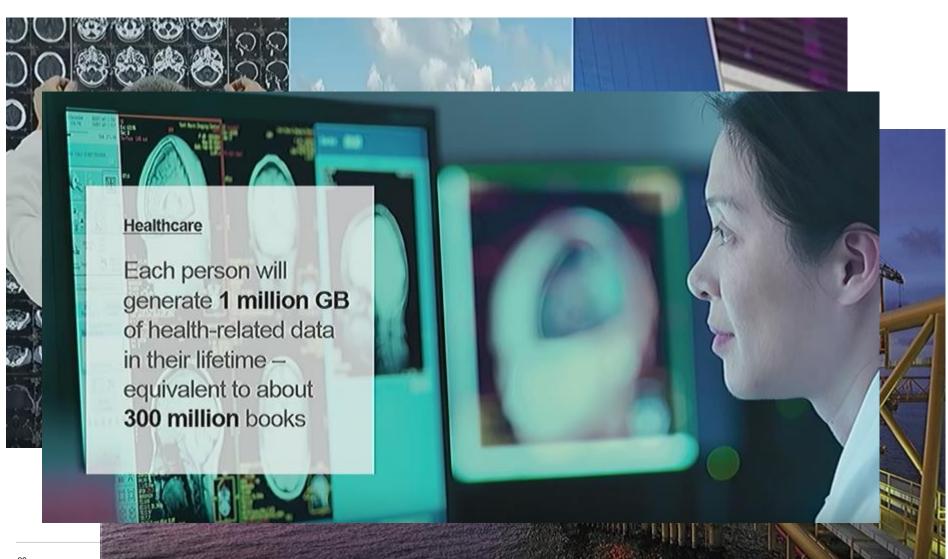
Watson is not miming or replicating the brain

- 80 90% of todays data is dark data not visible to computer systems
- Similar to dark material in space



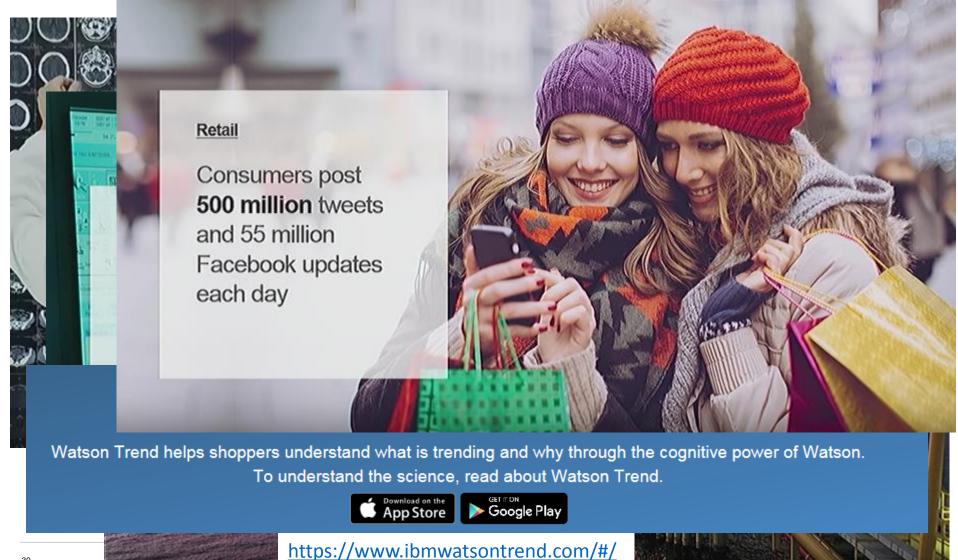
Watson is not miming or replicating the brain

- 80 90% of todays data is dark data not visible to computer systems
- Similar to dark material in space



Watson is not miming or replicating the brain

- 80 90% of todays data is dark data not visible to computer systems
- Similar to dark material in space





Cognitive in industry - predictions

CONSIDER:

Cognitive systems can understand the world through sensing and interaction, reason using hypotheses and arguments and learn from experts and through data. Watson is the most advanced such system.

Today, businesses in

36 countries across.

17 industries are applying cognitive technologies.

There are

350+

Watson ecosystem partner companies with

100 of those have taken their product to market.

78%

of business and IT execs believe successful business will manage employees alongside intelligent machines.

On average there are

1.3B

Watson API calls a month and growing.

Among C-Suite executives familiar with cognitive computing:



in **insurance** intend to invest in cognitive capabilities.



in **healthcare** believe it will play a disruptive role in the industry, and 60% believe they lack the skilled professionals and technical experience to achieve it.



in **retail** intend to invest in cognitive capabilities.



in **telecommunications** believe it will have a critical impact on the future of their business.



Cognitive Business

News room > News releases >

IBM Launches Industry's First Consulting Practice Dedicated to Cognitive Business

Specialists in analytics and data science lead client journeys spanning cognitive banking, retail, supply chain and healthcare, among others

Select a topic or year

◆ Related XML feeds

Armonk, N.Y. - 06 Oct 2015: IBM (NYSE: IBM) today launched the industry's first consulting organization dedicated to helping clients realize the transformative value of <u>cognitive</u> business.

IBM Cognitive Business Solutions extends the exclusive cognitive leadership of IBM <u>Watson</u> and the company's established market leadership in business <u>analytics</u>. The new practice draws on the expertise of more than 2,000 consulting professionals spanning machine learning, advanced analytics, data science and development, supported by industry and change management specialists to accelerate client journeys to cognitive business.

Cognitive represents an entirely new model of computing that includes a range of technology innovations in analytics, natural language processing and machine learning. Industry analyst firm IDC predicts that by 2018, half of all consumers will interact regularly with services based on cognitive computing.

"Our work with clients across many industries shows that cognitive computing is the path to the next great set of possibilities for business," said Bridget van Kralingen, senior vice president, IBM Global Business Services. "Clients know they are collecting and analyzing more data than ever before, but 80 percent of all the available data -- images, voice, literature, chemical formulas, social expressions -- remains out of reach for traditional computing systems. We're scaling expertise to close that gap and help our clients become cognitive banks, retailers, automakers, insurers or healthcare providers."

A survey of more than 5,000 C-suite executives to be released this fall by IBM's Institute for Business Value (IBV) finds that executives from the highest-performing companies place significantly greater priority on cognitive capabilities than peers in market-following enterprises. Industry-specific IBV research shows that:

https://www-03.ibm.com/press/us/en/pressrelease/47785.wss



Infrastructure tuned for the Cognitive Era

Integration



Convert to an API economy

With the industry's most robust API management integration platform, IBM helps companies innovate securely.

(P)

Watch the video (01:18)

Data



Profit from data in near real-time

Quintillions of data bits are generated every day. IBM systems help companies take advantage with analytics services.



Watch the video (01:17)

Operations



Anticipate the future

Earthquake-proof systems. Middleware that detects, automates and solves problems. Always-on operations.

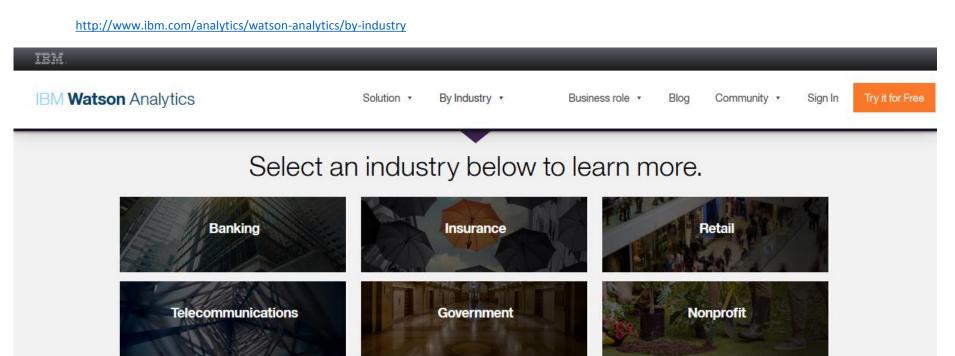


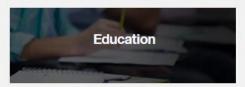
Watch the video (01:08)

IT Infrastructure for Cognitive

http://www.ibm.com/it-infrastructure/us-en/?lnk=buit

Experience Watson analytics –made easy







Bring your data to life. Create a free account here.

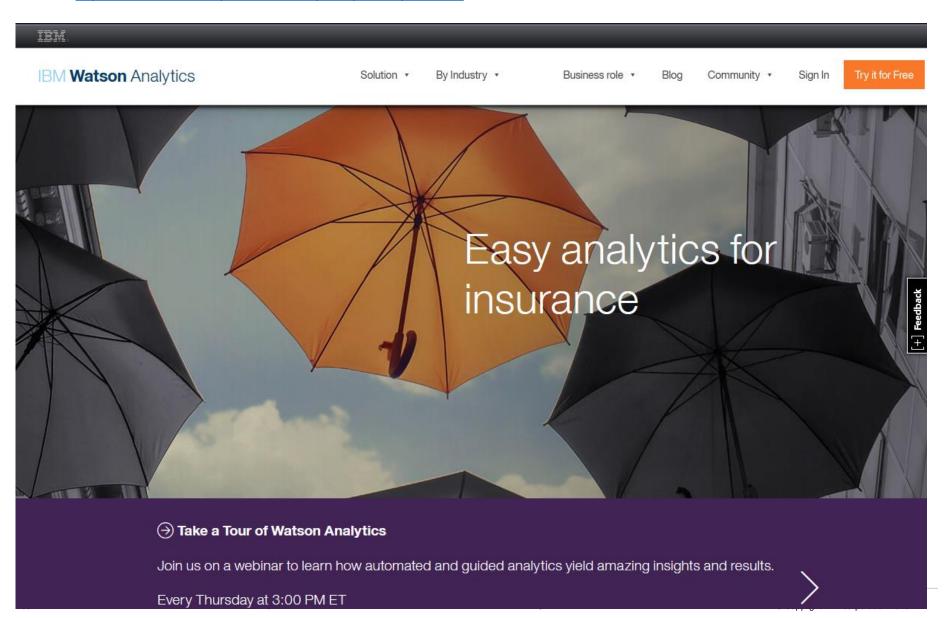




Deutsche Reg

Experience Watson analytics –made easy

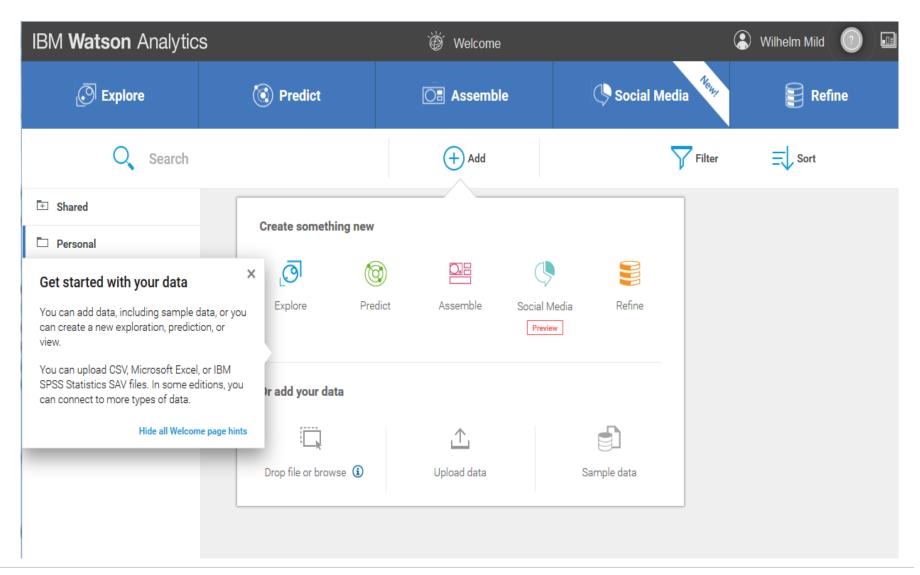
http://www.ibm.com/analytics/watson-analytics/by-industry/insurance





Experience Watson analytics

http://www.ibm.com/analytics/watson-analytics/by-industry



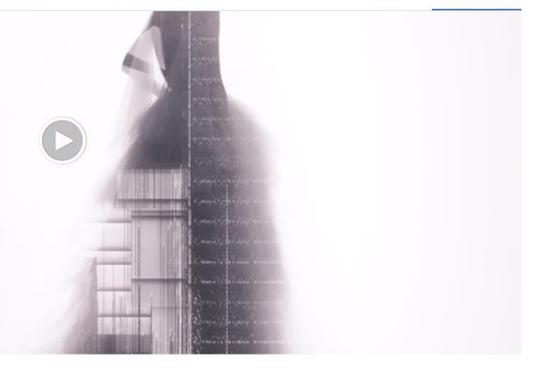


https://www.ibm.com/cognitive/outthink/stories/?iio=CCA&pid=17211371&cmp=usbrb&cm=b&csr=agus 2016 OUTHINK 1601 18&cr=time&ct=C34403XW&cn=Agenda outthink competitors 300x600 HTML5

Cloud at work Sports Cognitive at work

Marchesa + IBM Watson™ debut their first collaboration: the cognitive dress.

The masters of couture collaborated with a cognitive system to create a fashion first for this year's Met Gala: A dress that understands its fans, reasons with data and learns from hundreds of images.











outthink Cloud at work Sports Cognitive at work



"Using all available data to help you be a better athlete is the right thing to do."

- Tom Watson, Masters Tournament Champion, '77, '81











http://www.zdnet.com/article/ibm-and-friends-commit-to-linux-on-the-mainframe/

IBM z13 and LinuxONE – Fit for Cognitive computing

An Enterprise grade Linux solution portfolio

Data and Analytics

IBM InfoSphere
IBM DB2 BLU
SPARK

Cloud

Cloudant

Custom Patterns for Linux on z Systems

Trusted Computing

Spectrum Scale (GPFS technology)

IBM zAware V2.0

Crypto Express5S

GDPS Virtual appliance



Outstanding Capacity

IBM z13

Mobile

Node.js Internal Integration

Openness and Pricing

OpenSource and KVM

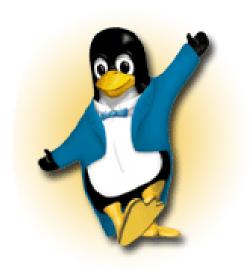
Large memory

High scalability

Security



Questions?



Dr. Manfred Gnirss

IBM Client Center Boeblingen, Germany

IBM Deutschland Research & Development GmbH Schönaicher Strasse 220 71032 Böblingen, Germany

Office: +49 (0)7031-16-4093 gnirss @de.ibm.com















Want to see and know more about Spark on z Systems? Then be our Guest for the Spark on z Systems Event



and analyzed by data scientists using their favorite languages and tools.

project in near future.

For more information click here Contact: Khadija Souissi, souissik@de.ibm.com