Data and AI
Get the most out of your data with IBM Cloud Pak for Data

Boris Korolkov
Hybrid Data Management leader,
IBM Cloud, Central & Eastern Europe
Boris.Korolkov@cz.ibm.com
1. Business Value
   Why is this relevant for our customers?

2. Solution Overview
   What exactly is Cloud Pak for Data & tangible benefits it offers to customers?

3. Use cases
   Key use-cases that we can position for Cloud Pak for Data

4. Key differentiators
   Capabilities that are unique to Cloud Pak for Data
Cloud Pak for Data

1. Business Value

**The Market**

*Digital transformation & AI is disrupting every enterprise & industry including our own*

**The Problem**

*Increasing data volumes & diversity; Growing regulations; Outdated monolithic systems*

**The Solution**

*Analytics platform with prescriptive information architecture & cloud agility behind the firewall*
Multi-Cloud & AI are driving Digital Transformation

### Business Value

- **75%**
  - of large enterprises will have digital transformation at the center of corporate strategy within two years

- **81%**
  - of companies do not yet understand the data required for AI

- **>85%**
  - of enterprise IT organizations will commit to multi-cloud architectures

- **>59%**
  - of large enterprises see improved app quality & reduced defects with containers
Barriers to successful AI transformations
There is no Artificial Intelligence (AI) without Information Architecture (IA)

Data Ecosystem
- Data in silos
- Difficult to access
- No lineage

Analytics Tools
- Discrete tools
- Different preferences
- Difficult to manage

Workflow
- Not integrated
- Not governed
- Lack dev/prod parity

Culture
- Not collaborative
- Slow provisioning
- Lack trust in AI

There is no AI, without IA
Cloud Pak for Data
Helps make your data ready for AI and Cloud

1. Business Value

2. Ladder to AI

   - **Organize** – Create a trusted analytics foundation
   - **Analyze** – Scale insights with ML everywhere
   - **Infuse** – Deploy trusted AI-driven business processes

**Strong Foundation** – Built on “Cloud native architecture”
Cloud Pak for Data—*Cloud Native by Design*

**Agility**

An architecture of loosely coupled data services, easily refactored to create containerized workloads

**Efficiency**

Stand-alone workloads composed of micro-services & data that are flexibly deployed, orchestrated and managed

**Cost Savings**

Agile provisioning of containerized workloads in multi-Cloud environments and consumption of Cloud services

---

**Cloud Pak for Data**

- **Microservices**
- **Containerized Workloads**
- **Multi-Cloud Provisioning**

**Supporting Technologies**

- Cloud Foundry
- Docker
- Kubernetes
- Helm
- Terraform
The building blocks of AI

Collect
- All Sources of Data
- The Common Application layer
- Write once, deploy anywhere

Organize
- Find, Catalog, mask data
- Built in compliance
- Advanced transformation capabilities

Analyze
- Open platform for Data Science
- Descriptive, predictive to prescriptive
- ML deployment

Some Assembly Required – products available
- Db2 & Db2 Warehouse
- Db2 Event Store
- Integrated Analytics System
- Big SQL & Hortonworks Hadoop
- Information Server
- Data Replication
- Master Data Management
- Optim & StoredIQ
- SPSS & DSX
- Cognos & Watson Analytics
- Watson Explorer
- Planning Analytics

2. Solution Overview
The building blocks of AI

- **Collect**
  - All Sources of Data
  - The Common Application layer
  - Write once, deploy anywhere

- **Organize**
  - Find, Catalog, mask data
  - Built in compliance
  - Advanced transformation capabilities

- **Analyze**
  - Open platform for Data Science
  - Descriptive, predictive to prescriptive
  - ML deployment
  - Analyze insights on demand

**Integrated end-to-end platform on**

**Cloud Pak for Data**
Collect Data

1. Provision, host & manage data sources on the platform

- Db2 Event Store
  - In-memory data store capable of extremely high speed ingest and deep, real-time analytics.

- Db2 Advanced Enterprise Server Edition
  - Relational database that delivers advanced data management and analytics capabilities for transactional and warehousing workloads.

- Db2 Warehouse
  - Data warehouse designed for high-performance, in-database analytics.

- IBM Db2 for z/OS
  - Create databases in Db2 for z/OS and work directly with the data from IBM Cloud Private for Data.

- Data Virtualization
  - Query across many data sources as one.

- MongoDB
  - Scalable, open source NoSQL database.

Sample Data Sources

- Db2
- HDFS
- Custom JDBC
- Cloudera
- Oracle
- Teradata
- Hortonworks
- Netezza

2. Connect to Existing data sources and/or upload your data

Add data source

- [Input source name]
  - Type the data source name here.

- [Description]
  - Type your description here.
1. **Auto-discover, catalog and publish meta data, track lineage etc.**
Organize Data

2. Solution Overview

3. Transform & Migrate Data – Build & Execute ETL jobs @ scale
Analyze Data

4. Build Visualizations & Dashboards

Product Line
- Camping Equipment
- Golf Equipment
- Mountaineering Equipment
- Outdoor Protection
- Personal Accessories

<table>
<thead>
<tr>
<th>Executive Summary</th>
<th>Geo Segmentation</th>
<th>Order Type Analysis</th>
<th>Margin Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Product Line</strong></td>
<td><strong>Camping Equipment</strong></td>
<td><strong>Revenue</strong></td>
<td><strong>Quantity</strong></td>
</tr>
<tr>
<td><strong>3.77B</strong></td>
<td><strong>69.1M</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Revenue</strong></td>
<td><strong>Quantity</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>0.41</strong></td>
<td><strong>Margin</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Revenue by Order Method
- Special
- Mail
- Fax
- E-mail
- Telephone
- Sales visits
- Web

Product Types (Size by Revenue)
- Safety
- Eyewear
- Watches
- Tents
- Cooking Gear
- Navigation
- Climbing Equipment
- Tools
- Packs
- Sleeping Bags
- Lanterns
- Camping Accessories

Revenue by Country
- [Map showing revenue distribution by country]
2. Solution Overview

5. Build, Deploy and manage Data Science & Machine Learning Models
6. A single Pane to administer & monitor your data & analytics workloads
Cloud Pak for Data- Architecture

Open, Extensible API Platform

- App Developers
- Business Partners

Personalized, Collaborative Team Platform

- Data Engineers
- Data Stewards
- Data Scientists
- Business Users

Instant, Pre-assembled Provisioning

- Collect Data
- Organize Data
- Analyze Data

Cloud-native Data Micro Services

- Databases on-demand
- Data warehousing
- Fast data ingest / Streaming data
- Data Virtualization
- Data integration and Transformation
- Data curation
- Governance, privacy: policies & rules
- Data asset lifecycle management
- Data visualization & exploration
- Machine learning & deep learning
- Model management & deployment
- Dashboards & business reporting

Admin & Ops Dashboards

Enterprise Data Catalogs

IBM Cloud Private

- Elastic Runtimes
- Product Ready Operations
- Built-in Continuous Delivery
- Secure to the Core
Cloud Pak for Data
Increases workforce productivity across the analytics lifecycle – governed seamlessly

Data Engineer
Architects data pipelines & ensures operability

Data Steward
Governs data and ensures regulatory compliance

Data Scientist
Gets deep into the data to draw insights for the business

Business Analyst
Works with data to apply insights to business strategy

App Developer
Plugs into analysis and code to build apps

Decision Maker(s)
CXO
CIO / CDO / CAO
LOB Leader
Sys Admin
Packaging & Client Adoption Architecture (Big Picture)

Analogy: iPhone Apps

Apps by Apple, 3rd Party sold through marketplace & custom

1. ICP for Data

Core Cloud-native Data platform Services with Configurable Installation Options

- Collect
- Organize
- Analyze
- Infuse

2. Add-Ons

Customize & Extend with add-on microservices

Add-On Services
- IBM
- Partner
- Custom

Cloud Platforms

Install on multi-cloud native platforms

Multi-Cloud
Cloud Pak for Data
Top 4 Use Cases

Manage your Data Anywhere

1. Manage all your enterprise data regardless of where it lives
   (Data Virtualization)

2. Gain control & leverage your data from connected devices
   (Fast data & Streaming analytics)

Operationalize Data Science & AI

Build, deploy, manage & govern models & data @ scale to improve business outcomes
(see following slides for details)

- a. Customer Churn
- b. Cross Sell / Up Sell
- c. Predictive Maintenance

Shift to Next-Gen workloads

Shift to Cloud Native

- a. Provision & scale in minutes
- b. Build once, deploy anywhere—multi cloud support
- c. Built in automation & collaboration to increase productivity

Smarter Governance

Governance to enable self service analytics

Auto-discover meta data, manage governance rules & policies, enforce privacy etc. to mitigate risk & ensure compliance

- e.g. accelerate GDPR Compliance
1. Open Ecosystem – *where we partner, co-create & lead*

4. Key Differentiators
Build Once, Run anywhere
- in your own data center or cloud infrastructure of your choice

Helps avoid Vendor Lock in
Multi-Cloud is Key to Organizational Agility

IBM is uniquely positioned today

- Portable Applications
- Private Cloud
- Public Cloud
- Manhattan

4. Key Differentiators

Cloud Pak for Data

New App Dev
Public Integration

Lift & Shift

Cloud Pak for Data
3. Data Virtualization

Value Proposition

– Query Anything, Anywhere - across data silos (heterogeneous data assets)

– Reduced costs with distributed parallel processing (vs centralized processing) - free of data movement, duplication, etc.

– Auto-discovery of source and metadata, for ease of viewing information across your organization

Key Differentiators

• Self-Discovering, Self-organizing cluster

• Integrates disparate data assets with simple automation, providing seamless access to data as one

• Governance, security & scalability by design
4. Operationalizing Data Science & AI

It's all about Day 2

4. Key Differentiators
5. Data Security

1. **Data Encryption** supported through storage layer (e.g. IBM Spectrum Scale)

2. **Built-in Authentication & Authorization**: Integrates with industry leading LDAPs (e.g. Microsoft Active Directory etc.) and supports SAML

3. **Data privacy & governance by design**: Data masking, Auto-discovery & curation

4. **Smarter Compliance**: Regulatory ML and other differentiating capabilities on roadmap.
Cloud Pak for Data – Ranked #1 by Forrester

Enterprise Insight Platforms - Definition

- Enterprise insight platforms pre-integrate most — or all — of the technology required to build systems of insight and thus help business move faster. The need to move faster and change more easily is the driving force behind customer demand for these platforms.

- Vendors that can better support all the personas of an insight team with unified experiences that feature governance and can creatively enable hybrid cloud and multi-cloud delivery will win.

Forrester’s commentary on “ICP for Data”

IBM has an impressive portfolio of individual data management and analytics capabilities that have consistently scored well on individual component Forrester Waves. With IBM Cloud Private for Data, IBM has pre-integrated capabilities that allow clients to be productive in a week or less. We were also impressed with its ML-assisted data cataloging and governance tools. IBM’s platform uses Kubernetes to deploy on-premises or into the public cloud. Lastly, IBM’s support for different insight team personas through tailored but unified experiences is commendable. Firms looking to unify the work of insight teams will do well on this platform.

Microsoft’s Perceived Weakness – Azure Cloud Platform

While Microsoft offers AI services, its multimodal predictive analytics and machine learning (PAML) tools scored poorly in previous Forrester Waves. Finally, we found this offering to be too light on data governance capabilities and self-service data preparation tooling, both of which are critical insight team capabilities.
IDC’s VP Ritu Jyoti:
In 2018 the AI market experienced:

- **steady growth 35.6%**
- revenue of **$28.1 billion**
- .. moving **from POC to full blown production**
- growth in all the technology categories from software to hardware to services...
Дякую за увагу!
Спасибо за внимание!
Thank you!

Boris Korolkov
Hybrid Data Management leader,
IBM Cloud,
Central & Eastern Europe
Boris.Korolkov@cz.ibm.com
Ibm.com
Дякую за увагу!