

Managing the Lifecycle in Big Data Environments

Wolfgang Epting
IBM Deutschland GmbH
Ehningen

Schlüsselworte

IBM InfoSphere Optim, ILM, Information Lifecycle Management, Big Data, Hadoop, Data Governance, Data Integration, Archivierung, Aufbewahrung, Revisionsicherheit, Anonymisierung, Datenschutz

Einleitung

Big Data Umgebungen ermöglichen es Unternehmen heutzutage, große Datenmengen, zu sammeln, zu verarbeiten und in Echtzeit zu analysieren. Neben den klassischen, finden sich dort viele neue, unstrukturierte bzw. semi-strukturierte, Datenformate wie zB. von Maschinen generierte Informationen, RFID Tags, Social Media Informationen etc.

In dieser Session werden bewährte Data Lifecycle Methodiken vorgestellt, um den Business Value von Data Warehouse und Big Data Umgebungen zu maximieren. Wir werden diskutieren, wie sich gemeinsame Regelungen über Systemgrenzen hinweg definieren lassen, in denen Aspekte des Datenschutzes, revisionssicherer Archivierungsverfahren, von Aufbewahrungsfristen und Lösungspraktiken geregelt und während des kompletten Lebenszyklus überwacht werden können.

Ein solcher vollumfänglicher Governance Ansatz muss alle Geschäftseinheiten und Funktionen berücksichtigen und darf sich nicht nur auf einzelne Bereiche beschränken. Nur dann kann das Thema Data Governance leicht auf das Design, die Verwaltung und die Erweiterung neuer Datenquellen übertragen werden.

Vortragstitel

Managing the Lifecycle in Big Data Environments

Acting on Insight Requires Confidence in Data

Make Informed Decisions



Take Bigger, Calculated Risks



Uncover competitive advantages



Identify new opportunities



Information Integration & Governance for Big Data

Automated Integration



Rapid, easy access to big data, wherever it resides

Visual Context



Easy categorization, indexing, discovery of big data to optimize its usage

Agile Governance



Definition and execution of governance appropriate to data value and intended use

Abb.1: Information and Integration & Governance for Big Data

InfoSphere Delivers Critical Confidence for Big Data Use Cases



Big Data Exploration

- ✓ Understand confidence
- ✓ Determine risk



Enhanced 360° View of the Customer

- ✓ Establish master record
- ✓ Extent to all sources



Security/Intelligence Extension

- ✓ Automatic data protection
- ✓ Mask sensitive information



Operations Analysis

- ✓ High volume data integration
- ✓ Automatic data protection

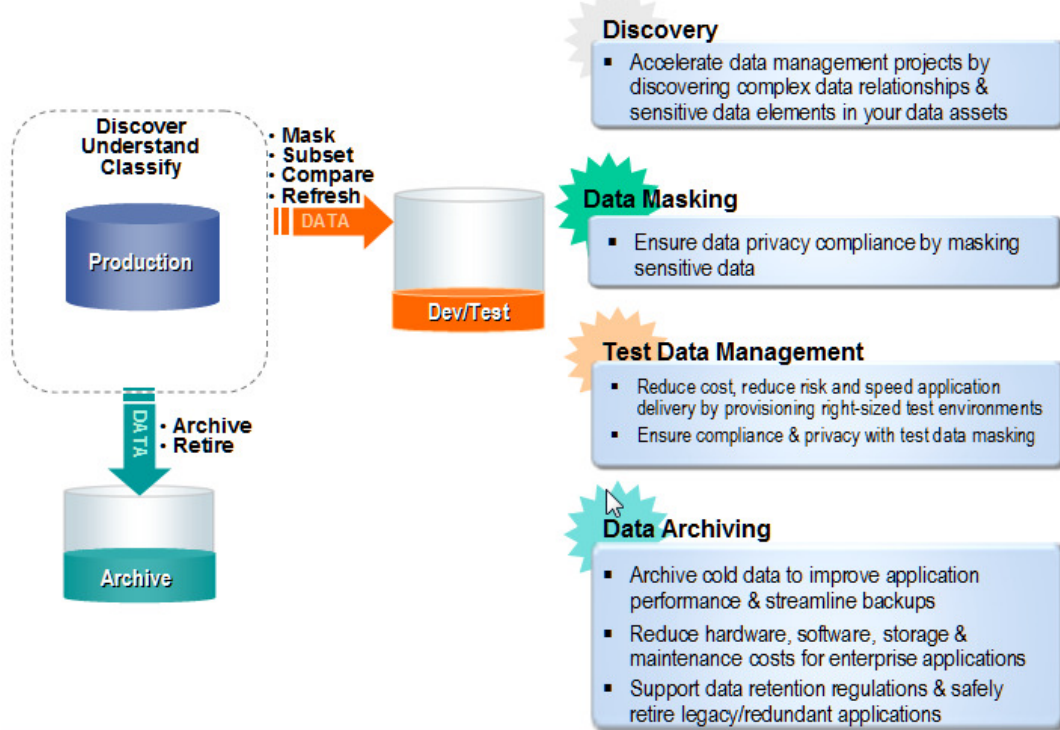


Data Warehouse Augmentation

- ✓ High volume data integration
- ✓ Agile big data archiving and retrieval

Abb.2: IBM InfoSphere delivers critical confidence for Big Data Use Cases

Overview: IBM InfoSphere Optim



Discovery

- Accelerate data management projects by discovering complex data relationships & sensitive data elements in your data assets

Data Masking

- Ensure data privacy compliance by masking sensitive data

Test Data Management

- Reduce cost, reduce risk and speed application delivery by provisioning right-sized test environments
- Ensure compliance & privacy with test data masking

Data Archiving

- Archive cold data to improve application performance & streamline backups
- Reduce hardware, software, storage & maintenance costs for enterprise applications
- Support data retention regulations & safely retire legacy/redundant applications

Abb.3: IBM InfoSphere Optim - Overview

IBM InfoSphere Optim solves key data challenges

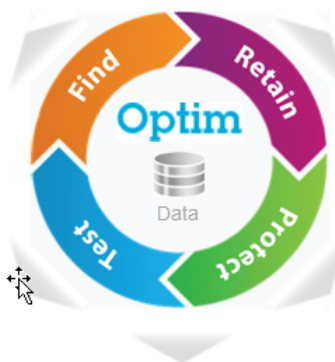
Identify Relevant & Sensitive Data
Find what data must be retained, protected or removed

Dispose of Unnecessary Data
Remove unnecessary data from critical transactional or analytics applications

Retain Essential Data
Historical inactive data is safely retained while easily accessible for reports and compliance

Optimize Test Data
Automate and optimize the application test processes that rely on data

Protect Sensitive Data
Private Data: Customer IDs, credit cards and financial data are masked or redacted



↑ Application Performance

↑ Availability

↓ Data Security Risk

↓ Costs

↑ Speed to make changes

Abb.3: IBM InfoSphere Optim – Solution of Key Data Challenges

Kontaktadresse:

Wolfgang Epting

IBM Deutschland GmbH

IBM-Allee 1

D-71139 Ehningen

Telefon: +49 (0) 160-9064 3048

E-Mail wolfgang.epting@de.ibm.com

Internet: www.ibm.com