

Inhaltsverzeichnis

Teil I	Grundlagen	1
1	Einführung in die BIA-Architekturen	3
	Peter Gluchowski • Frank Leisten • Gero Presser	
1.1	BIA-Trends und -Entwicklungen	3
1.2	Architekturkonzepte und -facetten	5
1.3	Datenbezogene Rahmenbedingungen	8
1.3.1	Datenstrategie	9
1.3.2	Data Valuation	11
1.3.3	Data Management	13
1.4	Anforderungen an eine ganzheitliche BIA-Architektur	17
1.5	Klassische Architekturen für BIA-Ökosysteme	19
2	Architekturen und Technologien für Data Lakes	23
	Carsten Dittmar • Peter Schulz	
2.1	Historie der dispositiven Datenplattformen	23
2.2	Das Data-Lake-Konzept	24
2.3	Architektur eines Data Lake	27
2.4	Datenarchitektur eines Data Lake	31
2.5	Technologien für einen Data Lake	32
2.6	Herausforderungen in der Umsetzung eines Data Lake	36
3	Datenzugriffsstrategien für Analytics bei beschränktem Datenquellenzugriff	39
	Michael Daum	
3.1	Ursachen von Einschränkungen auf Datenquellen	40
3.2	BIA-Anforderungen an Datenquellen	44

3.3	Datenstrategische Überlegungen	44
3.3.1	Trennung von Problemstellung und technischer Lösung	45
3.3.2	Skalierbarkeit	46
3.3.3	Cloud-Strategie und Datenschutz	46
3.3.4	Data Management	47
3.4	Lösungsansätze bei unterschiedlichen Einschränkungen	47
3.4.1	Technische Probleme der Connectivity	47
3.4.2	(Firmen-)»Politische« Themen und Lizenzen	48
3.4.3	Datenschutzanforderungen beim Zugriff	49
3.5	Entkoppeln von Systemen und Datenvirtualisierung	49
3.6	Abgrenzung und weiterführende Themen	51
4	Enterprise Application Integration: aktuelle Ansätze	53
	Martin Janssen	
4.1	Ein altbekanntes Thema vor immer neuen Herausforderungen	53
4.2	Der Unterschied zwischen Theorie und Praxis ist in der Praxis weit höher als in der Theorie	55
4.3	Die Zeit des ESB	57
4.4	Neue Anforderungen durch die Clouds	59
4.5	Drei aktuelle Lösungsansätze	60
4.5.1	iPaaS – Integration mittels Low Code und als Turnkey-Lösung	60
4.5.2	Kafka – der neue ESB?	65
4.5.3	Serverless Integration – alles in der Cloud	68
4.6	Fazit	70

Teil II Plattformen und Ökosysteme **73**

5	Cloud Data Platform für die Logistikbranche: eine Lösung auf Basis von AWS	75
	Christian Schneider • Gero Presser	
5.1	Herausforderung	76
5.2	Grundlegende Architektur	77
5.3	Technische Architektur mit AWS	78
5.4	Data Lake: AWS S3 und AWS Lake Formation	81

5.5	ETL und mehr: AWS Glue	82
5.6	Data Warehouse: AWS Redshift	83
5.7	Query Engine: AWS Redshift Spectrum	83
5.8	Visualisierung: AWS QuickSight	84
5.9	Flexibilität in der Architektur	85
5.10	Betrieb und Wartung	86
5.11	Ergebnis und Resümee	86
6	Organise the world's data – like Google	89
	Stefan Ebener • Stiv Sterjo • Sascha Kerbler • Andreas Ribbrock • Alex Osterloh • Diana Nanova • Christine Schulze • Lukas Grubwieser	
6.1	Einführung	89
6.1.1	Herausforderungen für eine erfolgreiche BI-Landschaft	91
6.1.2	Der Nutzen einer erfolgreichen BI-Landschaft	92
6.2	BI in der Public Cloud vs. On-Premises BI	93
6.2.1	Vom Budgetprozess hin zum aktiven Kostenmonitoring ...	94
6.2.2	Neue Unternehmensstrukturen rund um BIA in der Cloud	94
6.2.3	Trennung von Datenspeicherung und Rechenleistung	95
6.2.4	Elastizität und Skalierbarkeit	96
6.2.5	Infrastruktur als Code – IaC	97
6.2.6	Konvergenz von SQL und KI/ML	97
6.2.7	Vom Prototyp zur Applikation	98
6.3	Moderne Ansätze und neue Konzepte für BIA	98
6.3.1	Data Mesh aka Enterprise Data Evolution	98
6.3.2	Lake House als nächste Generation des Data Lake	102
6.4	Business Intelligence mit Google Cloud	106
6.4.1	Einführung einer serverlosen Architektur	108
6.4.2	Einführung innovativer KI/ML-Technologien	109
6.4.3	KI/ML im produktiven Einsatz	112
6.4.4	Ende-zu-Ende-Anwendung von einer mit KI/ML integrierten Datenplattform	116
6.4.5	Das »Big Picture« als Referenzarchitektur für ein modernes Lake House	116
6.4.6	Betrieb produktiver Anwendungen mit Google	118
6.5	Fazit und Ausblick	119

7	Die Modern-Data-Warehouse-Architektur von Microsoft	121
	Fabian Jogschies	
7.1	Datenablage mit Azure Data Lake Storage Gen2	121
7.2	Data Ingest und Orchestrierung	125
7.3	Transformation, Serving und ML mit Azure Synapse Analytics	127
7.4	Transformation, Serving und ML mit Azure Databricks	132
7.5	Data Lab Toolbox – Machine Learning	135
	7.5.1 Azure Machine Learning (AML)	135
	7.5.2 Azure Cognitive Services	138
7.6	Visualisierung mit Power BI	140
7.7	Data Governance mit Azure Purview	141
7.8	Azure DevOps	142
8	SAP Business Warehouse von gestern bis morgen	145
	Daniel Eiduzzis	
8.1	Business Intelligence made in Walldorf	145
	8.1.1 SAP Business Warehouse – Wie alles begann	145
	8.1.2 Probleme, Kritik und Herausforderungen im SAP BW-Kontext	147
8.2	Entwicklung des SAP BW	148
	8.2.1 Die Zeit vergeht, das SAP BW bleibt	148
	8.2.2 Der große Wurf bleibt aus	150
8.3	SAP Business Intelligence – heute und morgen	152
	8.3.1 HANA und Cloud geben die Strategie vor	152
	8.3.2 Features und Werkzeuge für Data Management und Data Integration	155
	8.3.3 Reporting und Analyse dort, wo die Daten generiert werden	156
	8.3.4 Hybride Konzepte als State-of-the-Art-Architektur	156
8.4	Ausblick und Fazit	158

9	Aus der Theorie in die Praxis – der Einfluss regulatorischer Anforderungen auf eine moderne Referenzarchitektur	159
	Thomas Müller • Lisa Anne Schiborr • Stefan Seyfert	
9.1	Aktuelle Herausforderungen	159
9.2	Historisierung	163
9.2.1	Bitemporale Historisierung	164
9.2.2	Best Practice	165
9.3	Datenschichtenarchitektur	166
9.3.1	Datenschichten der Referenzarchitektur	166
9.3.2	Archivierung und Housekeeping	169
9.4	Integrationsarchitektur	171
9.4.1	Verfahren und Werkzeuge	172
9.4.2	Anbindung Metadatenmanagement	173
9.4.3	Anbindung Datenqualitätsmanagement	174
9.5	Metadatenmanagement (MDM)	174
9.5.1	MDM – Kernanforderungen	176
9.5.2	Die Metamodelllandkarte (Modellsichten)	178
9.5.3	Data Lineage	179
9.5.4	Best Practice MDM – Technologie	181
9.5.5	Best Practice MDM – Architektur	182
9.6	Datenqualitätsmanagement (DQM)	183
9.6.1	DQM – Kernanforderungen	184
9.6.2	Prüfregeln	184
9.6.3	Korrekturen	185
9.6.4	Best Practice DQM – Architektur	186
9.7	Fazit und Handlungsempfehlungen	187
10	Case Study: Crédit Agricole Consumer Finance Netherlands	191
	Nick Golovin • Don Seur	
10.1	Herausforderungen	191
10.1.1	Lange Time-to-Market	192
10.1.2	Zugang zu Echtzeitdaten	193
10.1.3	Daten für operative Zwecke	193
10.1.4	DSGVO-Konformität	193
10.1.5	Anbindung von modernen Datenquellen	193

10.2	Anforderungen an die neue Lösung	194
10.3	Moderne Datenarchitektur	195
10.4	Use Cases	197
10.4.1	360°-Blick auf Kunden	198
10.4.2	Echtzeit-Sales-Monitoring	198
10.4.3	Marketinganalysen	199
10.4.4	Risk Management	199
10.4.5	Data Preparation für regulatorische Reportings	199
10.5	Schlussfolgerung: Datenvirtualisierung das Allheilmittel?	200
11	Datenvirtualisierung	201
	Daniel Rapp • Thomas Niewel • Jörg Meiners	
11.1	Moderne Datenarchitekturen für das Zeitalter der Digitalisierung	201
11.2	Datenvirtualisierung – ein Überblick	202
11.2.1	Anwendungsfälle der Datenvirtualisierung	204
11.3	Die Technologie der Datenvirtualisierung	206
11.3.1	Zugriff auf das Datenmodell	207
11.3.2	Datenschutz und Sicherheit	208
11.3.3	Query-Optimierung	209
11.3.4	Daten-Caching	210
11.3.5	Datenkatalog	211
11.4	Abgrenzung zu anderen Integrationstechnologien	211
11.5	Kundenbeispiel: Die Festo Gruppe	212
11.5.1	Unternehmensprofil der Festo Gruppe	212
11.5.2	Geschäftsanforderungen	212
11.5.3	Die Lösung	213
11.5.4	Die Mehrwerte	214
11.6	Zusammenfassung	215
11.6.1	Die Anwenderperspektive	216
11.6.2	Die Data-Governance-Perspektive	216
11.6.3	Die IT-Perspektive	216

Teil III	Architekturbeispiele	217
12	BIA-Architekturen für klinische Studien	219
	Jörg Krempien • Jörg Frank • Philipp Kazzer	
12.1	Über klinische Studien	219
12.2	Anforderungen an die BI-Architektur	221
12.3	Architekturdetails	222
12.3.1	Architekturüberblick	222
12.3.2	Staging	223
12.3.3	Core	225
12.3.4	Publish	226
12.3.5	Virtualisierung (Domänen durch Konfiguration)	227
12.4	Entwicklungsgeschichte und Ausblick	229
12.5	Use Cases	230
12.5.1	Virtual Cut Off	230
12.5.2	Subject Status	231
12.5.3	Baseline Flags	231
12.5.4	Clean Patient Tracker	232
12.5.5	Fraud Detection in Clinical Trials	232
12.5.6	Testautomatisierung (TAT)	232
13	BIA-Architekturen in der Versicherungsbranche	235
	Gerhard Brückl • Timo Klerx	
13.1	Ausgangssituation	235
13.2	Zielsetzung	236
13.3	Zielarchitektur	237
13.4	Data Lake	241
13.5	Datenverarbeitung	245
13.6	Ablaufsteuerung	247
13.7	Data-Science-Labor	249
13.8	Reporting	250

14	BIA-Architekturen für kleine und mittlere Unternehmen	253
	Markus Begerow	
14.1	Ausgangssituation	253
14.2	Neue Themen als Treiber für Innovationen	254
	14.2.1 Stammdaten- und Datenqualitätsmanagement	254
	14.2.2 Cloud-Infrastrukturen	255
	14.2.3 Data Science im Mittelstand	260
14.3	Konsequenzen für kleine und mittlere Unternehmen	262
	14.3.1 Tabellen- und Textdateien ersetzen keine Datenbank	262
	14.3.2 Cloud-Servicemodelle verstehen	263
	14.3.3 Data Science light einführen	268
14.4	Fazit	272
15	Integrierte Planung und Reporting im Business-Analytics-gestützten Controlling	275
	Christian Fürstenberg • Oliver Zimmer • Björn Beuter	
15.1	Die Entwicklung der Finanzplanung und -analyse	276
15.2	Der Weg zur datengetriebenen Unternehmenssteuerung	278
15.3	CCH® Tagetik – eine Lösung für alle Corporate-Performance-Management-Bereiche	280
15.4	Data Literacy – Aufbau von Datenkompetenz im Controlling	282
15.5	Einsatz und Nutzen von Self-Service im Controlling	284
15.6	Power BI als Self-Service-Reporting- und Analyse-Architektur	286
15.7	Einsatz von Power BI im Umfeld von CCH®Tagetik	287
	Anhang	291
A	Autoren	293
B	Abkürzungen	305
C	Literaturverzeichnis	311
	Index	319