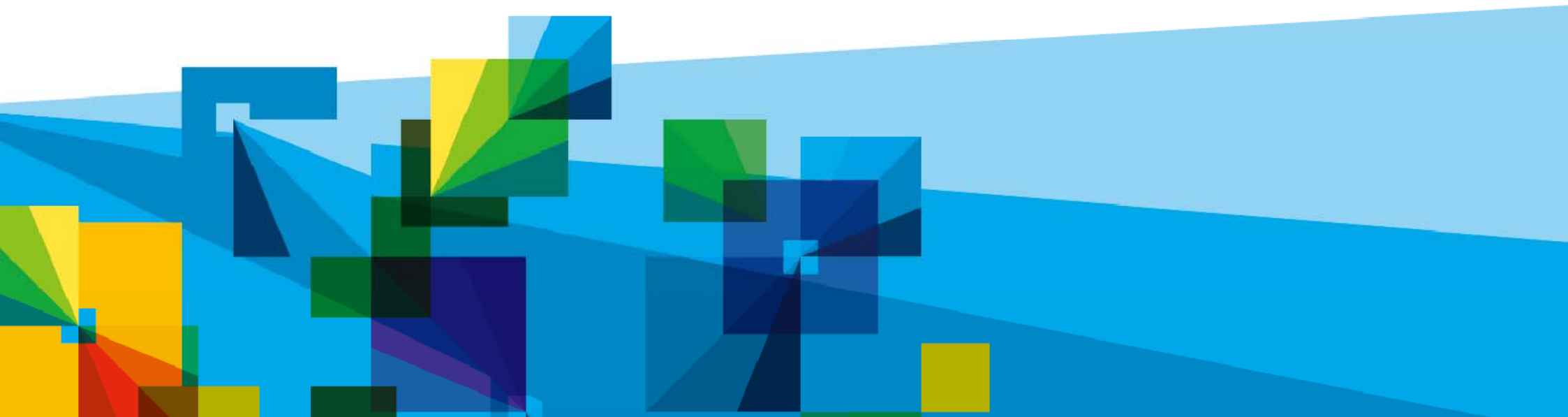




Big Data, Integration & Governance

28 - 30 August | Canberra; Melbourne; Sydney



Agenda



1

Big Data and MDM Part 1

2

What is Master Data Management?

3

What About Governance?

4

Big Data and MDM Part 2



So what about Big Data and Master Data Management

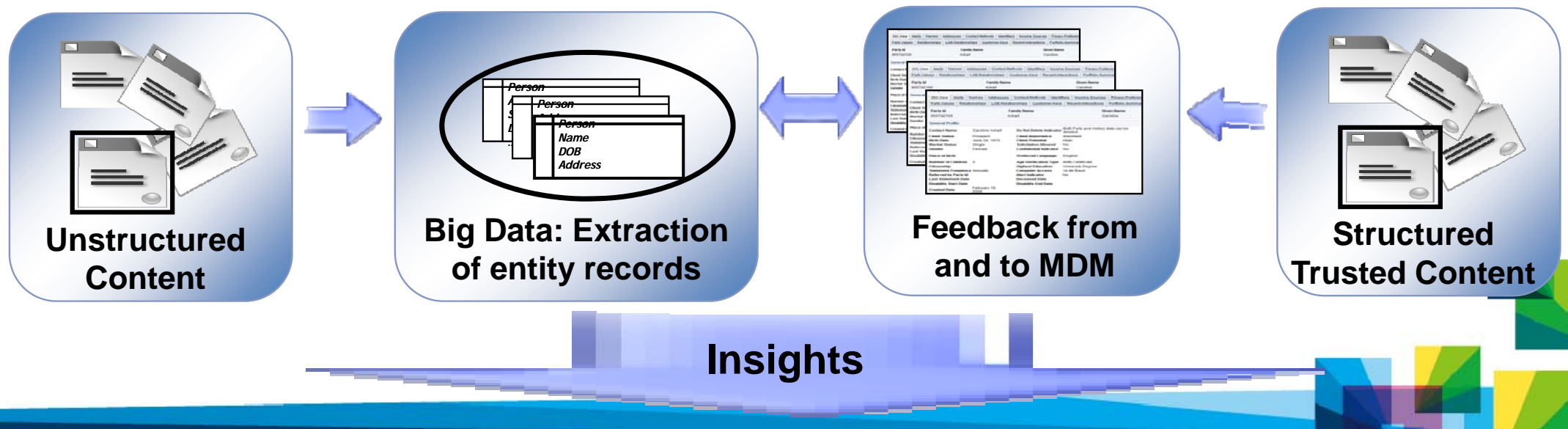
How do you make untrusted data trusted in real-time?

How do you discover relationships of people, places and things in big data and leverage them for your organisations benefit?



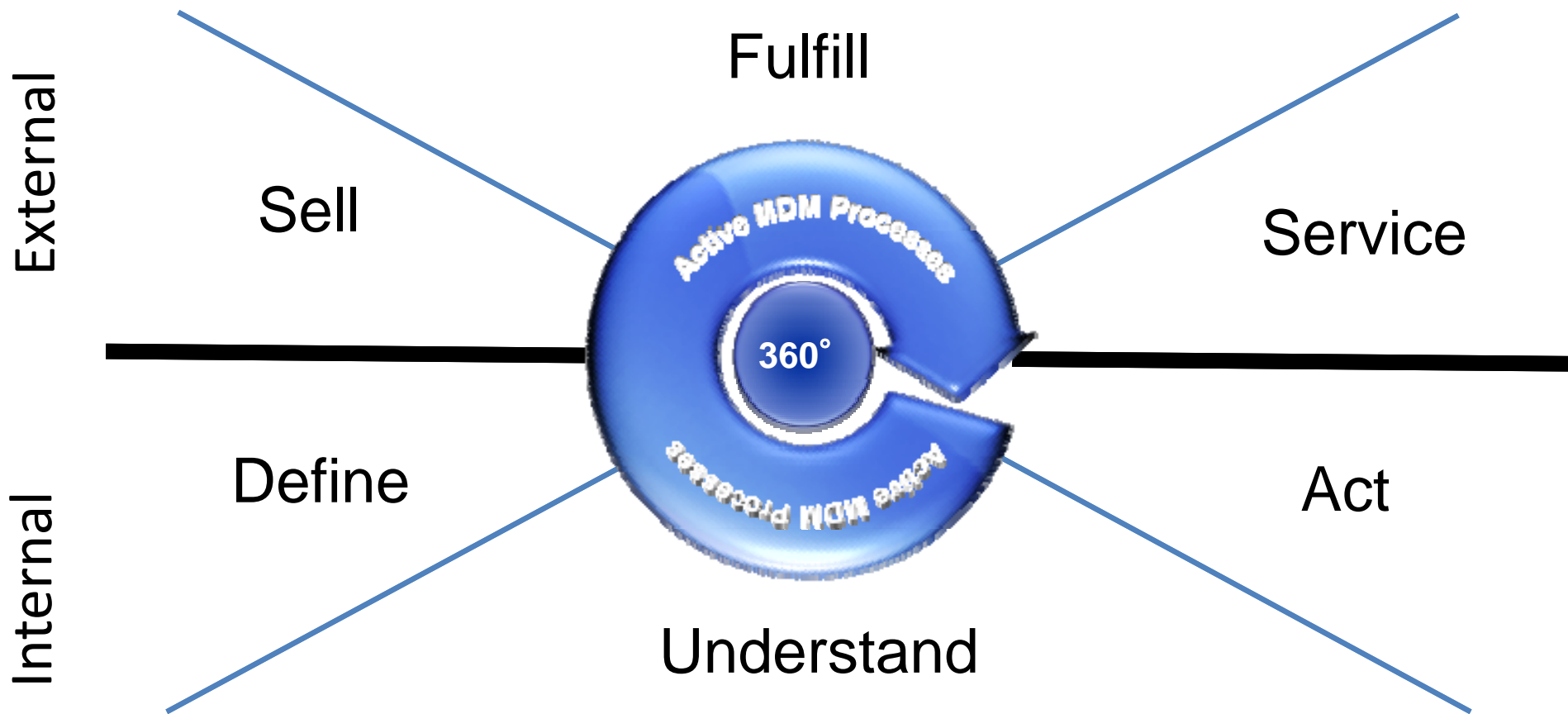
Big Data and Master Data Management

- Master Data Management bridges the gap between internal trusted data and external untrusted data
- High performance real-time matching of trusted information is the key to success.





What organisations do



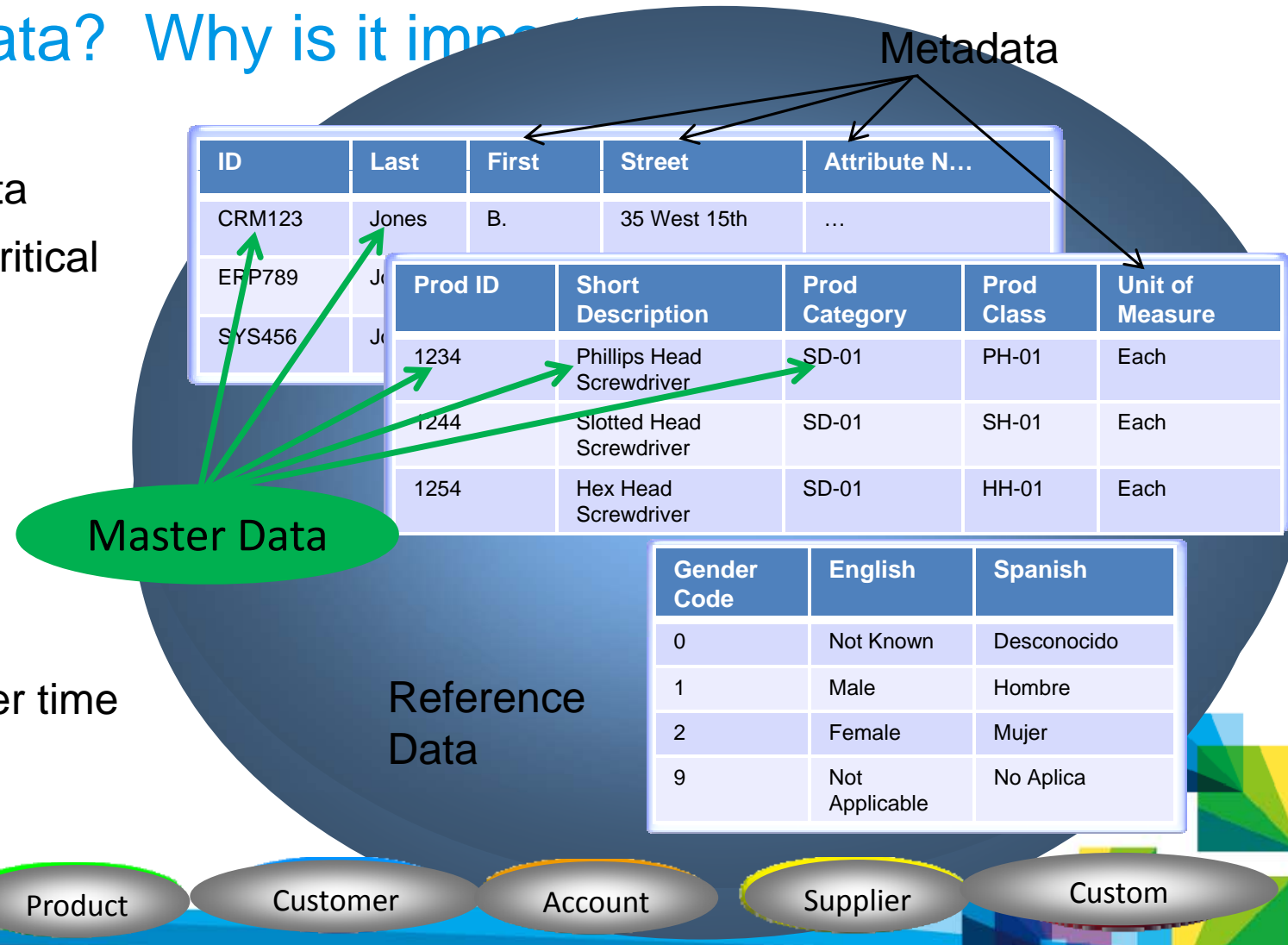
What is Master Data? Why is it important?

What is it?

- A subset of all enterprise data
- High-value, information for critical business processes
- Trusted data

Why is it important?

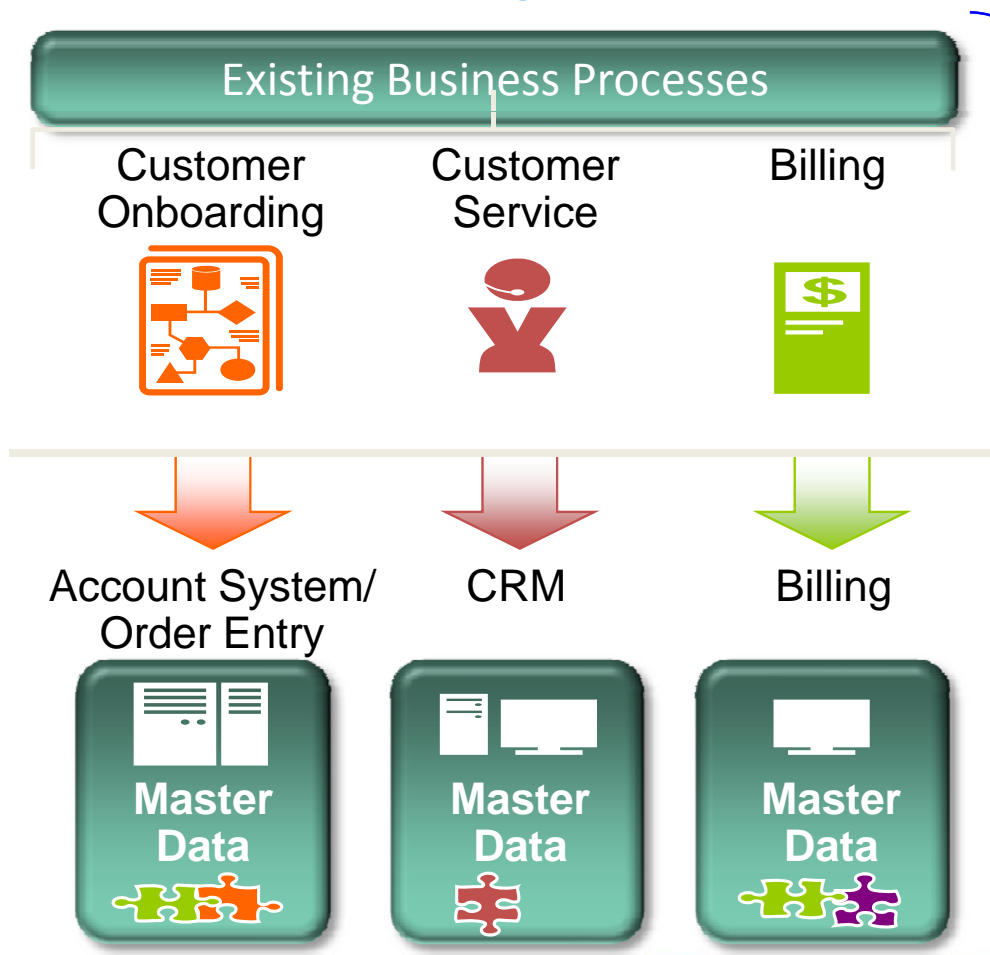
- Core to every transaction, application and decision
- Quality of data degrades over time and negatively impacts key business processes



The Issue – Fragmented master data

Business Process

Application



Data is fragmented among applications

Tightly coupled spaghetti

Who can you trust?

There is no process to manage data across applications

Change becomes hard



Master Data – Consistent Process

	CRM	FINANCE	B&M	ONLINE	WH
PARTY					
LOCATION					
PRODUCT					

- Ability to Implement new processes
- Multi-Channel Integration
- Compliance
- Cross- & Up-sell
- New Product Introduction
- Information Sharing



Information Governance creates order out of information chaos

“Data Governance is a control discipline for accessing, managing, monitoring, maintaining and protecting organisational information” *IBM Data Governance Council*

People



Process



Technology



The convergence of data quality, data management, data policies, business process management and management.

- MDM is the technology that implements the governance
- Governance is a continuous process, implement, monitor and improve.

Can you really do an Big Data or MDM project without doing governance?



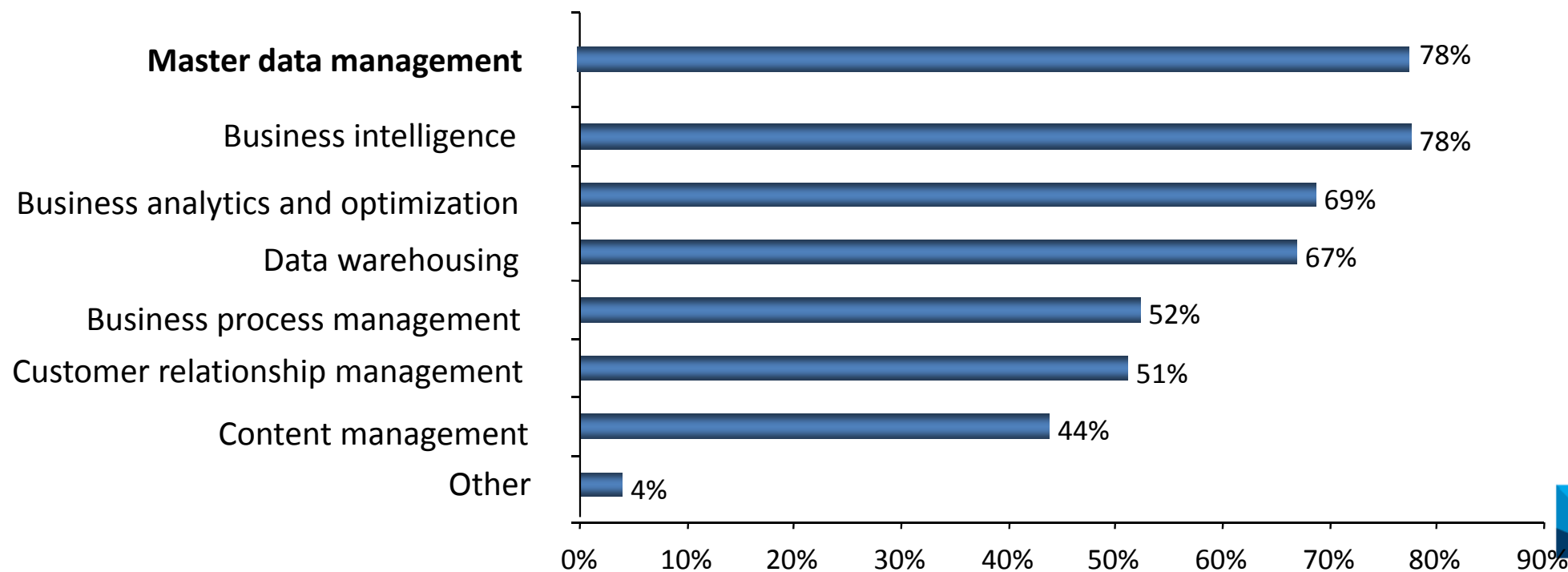
Ready, Code, Aim



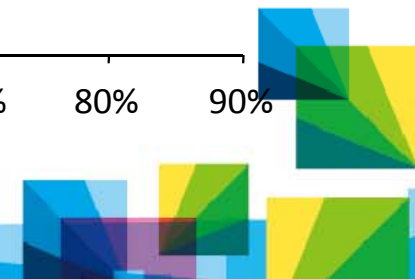


Information Governance makes IT Projects Successful

Information-related projects that require Information Governance



Source: BeyeNETWORK Custom Research Report Prepared for IBM Corporation



Information Governance increases MDM's success

- Many start without governance ...
- Many don't call it governance ...
 - ... but governance is *part* of any successful MDM program
- A Simple Test
 - What should happen when a customer changes their address at the **enterprise** level?
 - If you can't answer it – you need governance
- Governance enables MDM
 - It's the "How" in the story
 - It forces a conversation among the business constituents – the business needs to own it.





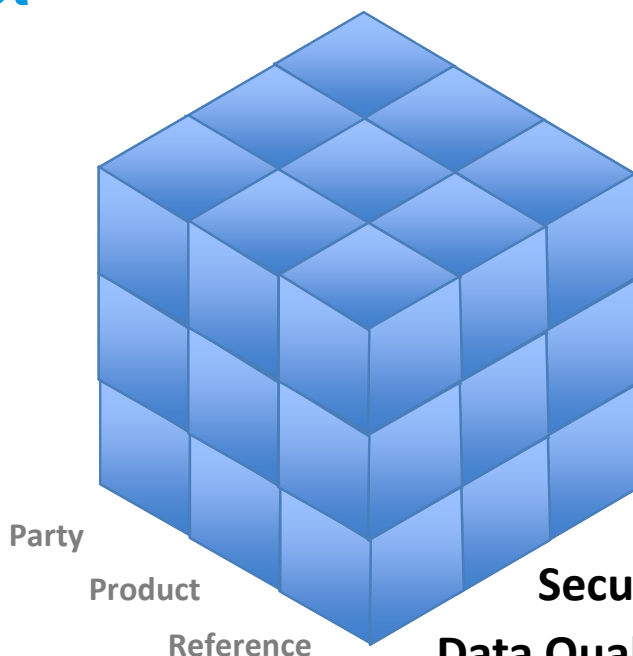
Master Data Governance Framework

Governance Aspect

MDM Governs **Master** Data, not **All** data

Policy Management

- Administration
- Enforcement
- Monitoring



Governance Aspect

Master Data Management

Metadata, Lifecycle Management

Security & Privacy

Data Quality





What is different about the governance of Master Data?

- Master Data Governance controls must be understood by IT and business.
- Focus is the key.
- Intertwined with enterprise data governance
- Active data governance
- MDM introduces two quality processes that require data governance control:
 - Creation and maintenance of the Golden/Trusted record(s)
 - Golden record proliferation to the enterprise (a complex process that typically can't be realised just by “feeding” the data to operational systems)





What is Different about the Governance of Master Data?

- MDM concepts such as matching and record linking, entity resolution, relationship resolution, attribute survivorship rules and the underlying data stewardship.
- Specific data quality metrics (KPI's) for data completeness, source-to-golden record consistency, accuracy, uniqueness.
- Access, security, audit, and privacy requirements (especially for party domain)
- Lineage requirements



Foundational vs. Business Specific Master Data Governance

- Must comply with business custom defined reference data codes
- If age < 18, must have guardian
- Product Catalog Rules

Custom - business process or data governance solutions (*Business Process Management, Business Rules*)

- Self-score completeness of the golden record
- Golden record uniqueness
- Source-to-golden record consistency
- Attribute accuracy, profiling

- Risk Management
- Basel II, Basel III, FATCA, APRA, internal Audit.

Industry/function specific - MDM-empowered applications: Customer on-boarding and global account opening, prospect referral, active data governance

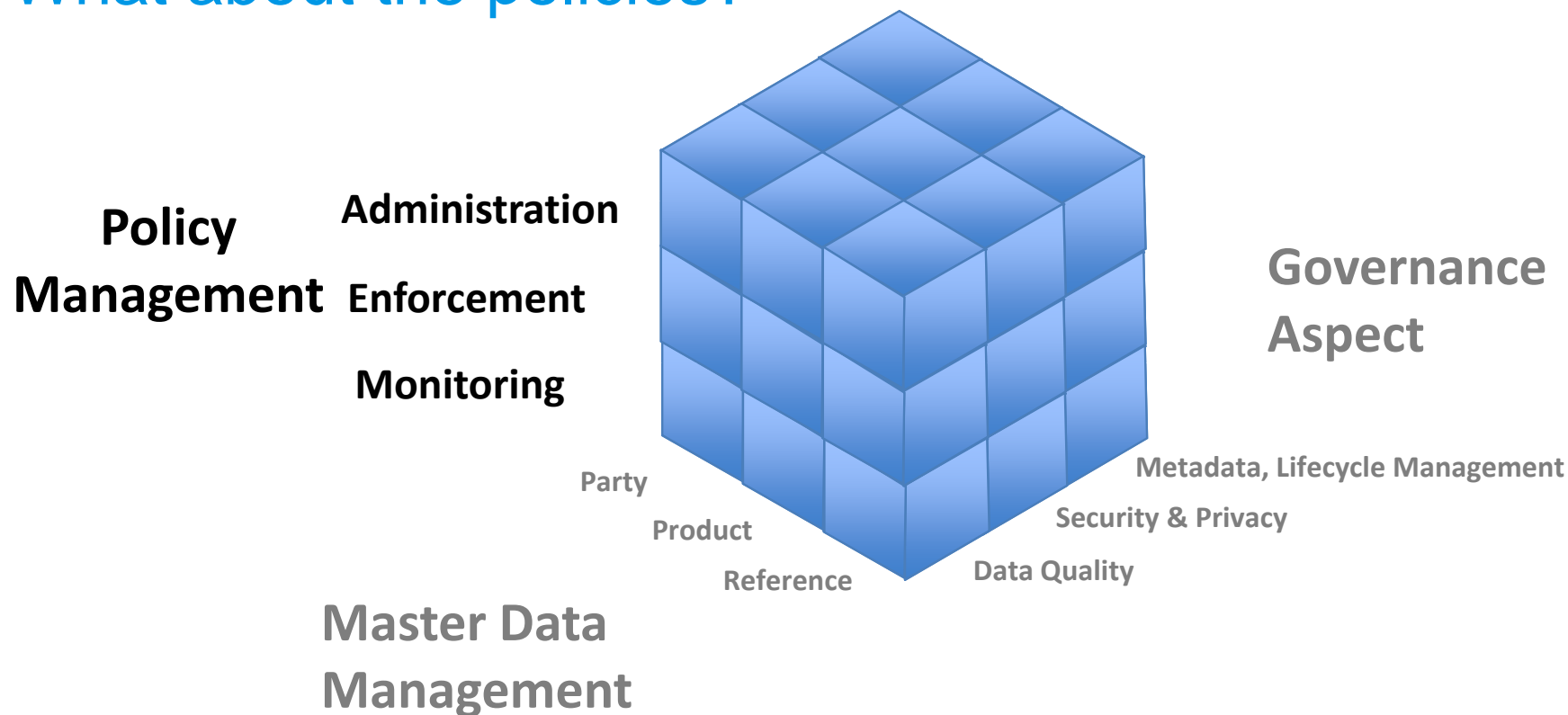
Foundational/generic MDG - data quality, stewardship, governance, policy management and security standards and controls that apply to a broad range of industries and companies (*Best Practices Layer; components are available OOTB*)



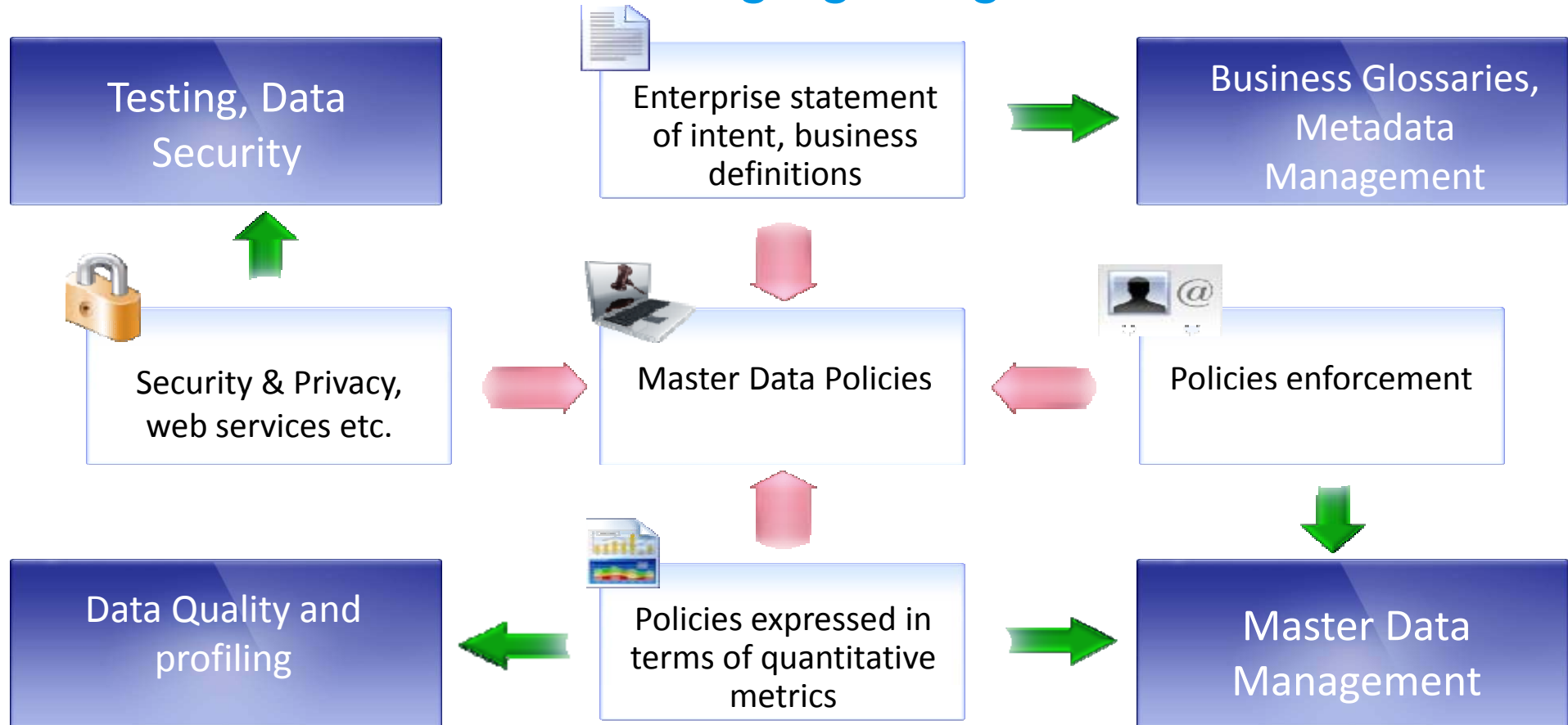


Master Data Governance Framework

What about the policies?



Master Data Policies – Bringing it together



Master Data Policy Management: Trust now and in the future

Policy Administration

- Address or Account Number invalid
- If < 18 years old, must have a guardian
- Self-score completeness < 170
- Source-to-golden record consistency < 80%
- Reference Code Unknown

Fine-grain metrics/KPIs and MDM-specific methods evaluation of DQ

Policy Enforcement

- “Hard” – when a record can’t be created by end user
- “Soft” – when the end user can be created (with a warning of policy violation); a policy violation task is created and sent to Inbox
- A “Soft” policy can be used for record entry thru ETL or messaging
- Release policy enforcement – the record can’t be released to the enterprise as it violates a release policy on record completeness
- BPM can be used for policy escalation/approvals

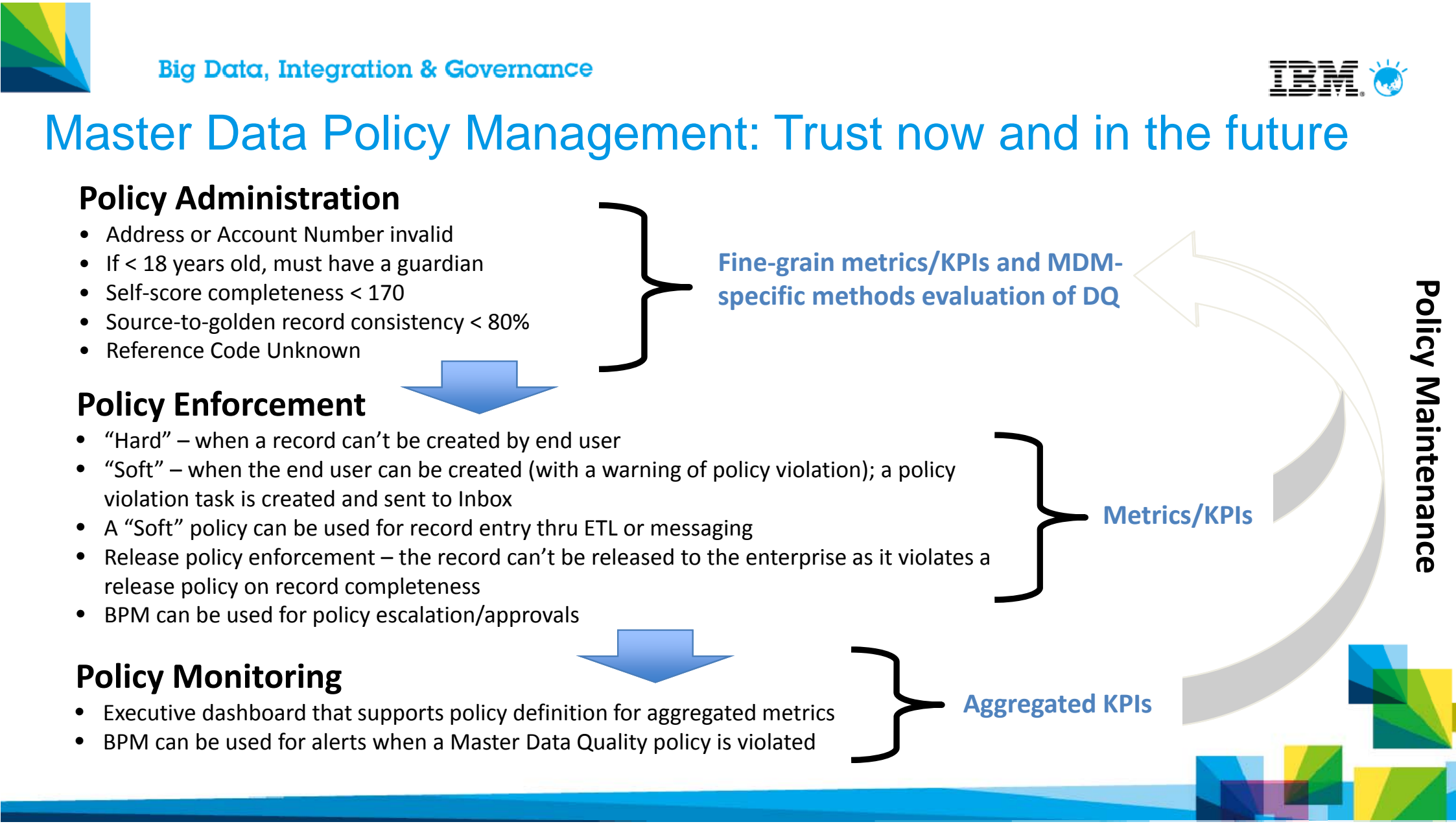
Metrics/KPIs

Policy Monitoring

- Executive dashboard that supports policy definition for aggregated metrics
- BPM can be used for alerts when a Master Data Quality policy is violated

Aggregated KPIs

Policy Maintenance

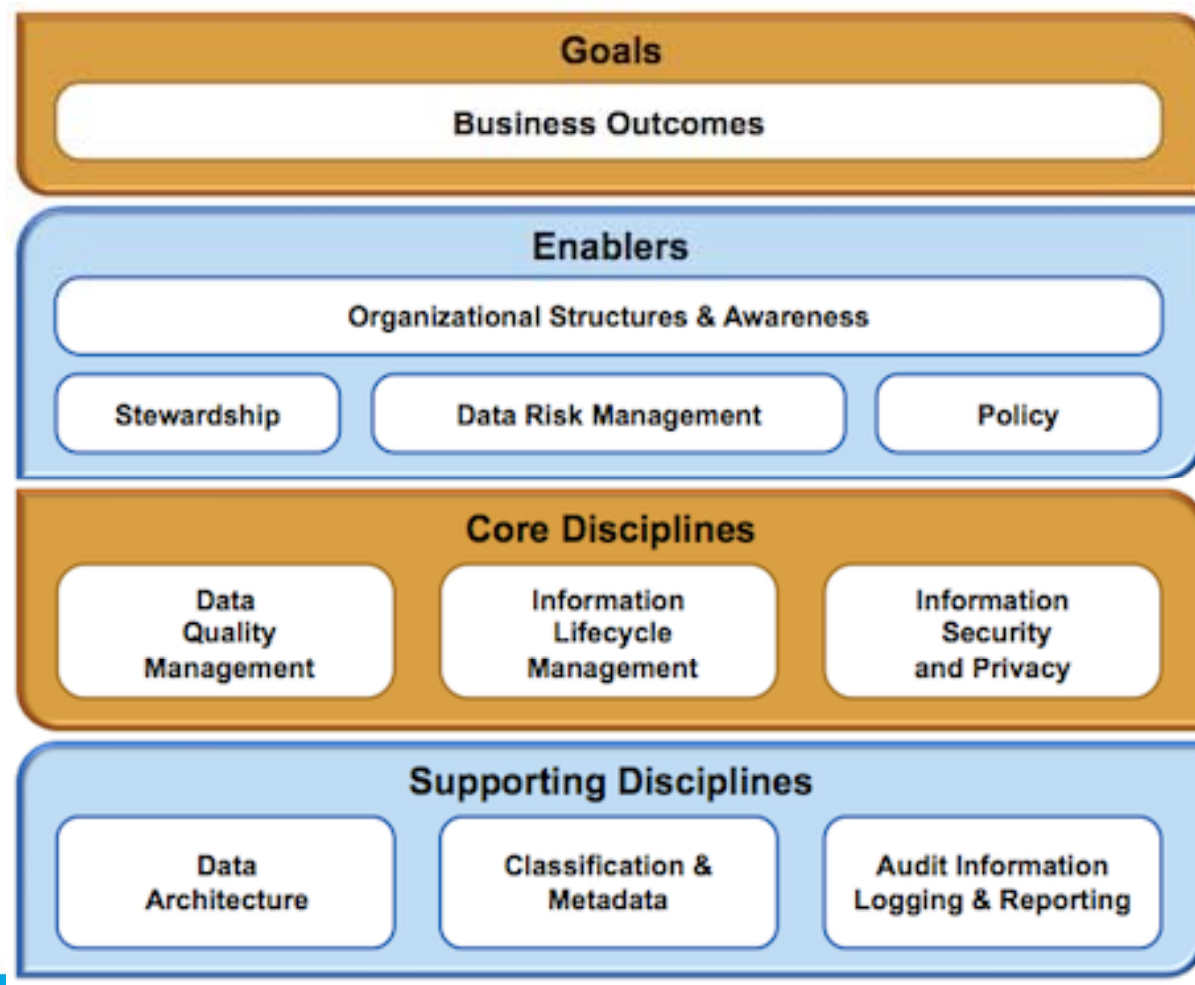




IBM Big Data Governance Framework

A Governance Framework to Assess Maturity

- Goals
- Enablers
- Core Disciplines
- Support Disciplines





Big Data Governance – Goals

- **Business outcomes**
 - Have you identified the key business stakeholders for the big data governance program
 - Marketing for social media governance
 - Legal for data retention policies
 - Have you quantified the financial benefits from big data governance
 - Reduced risk of fines and lawsuits due to data breaches
 - Avoiding negative impact on the brand due to bad publicity about the misuse of data
 - Public safety
 - Increased cross-sell and up-sell opportunities due to integration of social media with the master data environment






Big Data Governance – Enablers

- **Organisational structures and awareness**
 - Do you have a defined scope for big data that applies to your organization
 - Have you prioritized the types of big data that need to be governed?
 - Have any big data issues been addressed by the information governance council?
- **Stewardship**
 - How will you address the stewardship of big data?
 - Will the data stewards be responsible for gathering input from legal, marketing, and other departments regarding the acceptable use of big data
 - Are stewards empowered to define policies for retention of big data based on regulatory requirements and business imperatives?



Big Data Governance – Enablers

- **Data risk management**
 - Is risk management a key stakeholder for big data governance?
 - Have you established the linkage between big data governance and risk management?
 - Is there operationally realistic business continuity planning (technologies like Hadoop were not designed with traditional enterprise disaster recovery considerations)?
 - **Policy**
 - Have you documented a set of policies for big data governance?
 - Can these policies be inspected for enforcement?
 - Are you monitoring adherence to these operational controls using a Governance, Risk, and Compliance (GRC) framework?
- 



Big Data Governance – Core Disciplines

- **Data quality management**
 - Are your organisational policies on data quality being applied to both real-time (streaming) and at-rest (Hadoop) technologies?
 - Do you have consensus on the quality issues associated with big data where the value of the data may or may not be high or obvious?
 - Are you measuring quality?
- **Information lifecycle management**
 - What is the volume of storage for big data? What is the anticipated annual rate of growth?
 - Do you understand the business requirements that drive the retention of big data?
- **Information security and privacy**
 - Is the chief information security officer a key sponsor of the big data governance program?
 - Do you understand privacy regulations that affect big data (especially social media) by country, state, and province?





Big Data Governance – Supporting Disciplines

- **Data architecture**
 - Have you determined which applications should move in or stay out of the big data infrastructure platform?
 - Are you considering the impact of master data on big data?
 - How do you handle the lineage of big data?
- **Classification and metadata**
 - Does your organization-wide business terminology (business glossary) include key business terms relating to big data
 - Are you capturing key operational metadata to identify situations when big data is not loaded?
- **Audit information logging and reporting**
 - Do you have database administrators, contractors, and other third parties who possess unencrypted access to sensitive big data
 - Do you have a view of data Lineage?





What about the that Big Data project?

CRM FINANCE B&M ONLINE WH Big Data

	CRM	FINANCE	B&M ONLINE	WH	Big Data	
PARTY						
LOCATION						
PRODUCT						





Aiden
see's a
great video
on
YouTube



How do you establish trusted Big Data ?

The screenshot shows a YouTube video page. At the top, the search bar contains 'darth vader orchestra laroche'. The video title is 'Lego Star Wars - For the millionth time, I didn't make this' by user 'hidekatsu'. The video player shows a scene with a character holding a red lightsaber. Below the player, the view count is 27,255,923 and the comment count is 25,000. The description reads: 'A cute Lego Star Wars animation I'm sharing and yes, I didn't make this.' The video was uploaded on Nov 26, 2005, and has 39,920 likes and 2,855 dislikes. The artist is John Williams. The right sidebar shows a list of recommended videos, including 'Ab Workout by Flex Belt', 'YouTube Mix for John Williams', 'Star Wars Lego - Jedi vs Sith', 'Lego Star Wars 3 Commercial', 'Lego Star Wars 3 the Clone Wars Hostage', and '2010 LEGO STAR WARS - Emperor Palpatine's'.

9 other videos

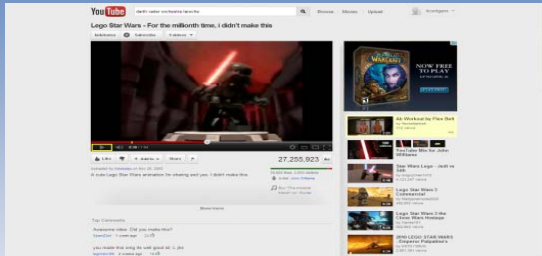
27 Million Views

25,000 Comments



Discovery of Intent to Purchase

You Tube



Top Comments

I want to buy DARTH VADER LEGO
[Aiden22](#) week ago 24 👍

Aiden



Dave



MDM

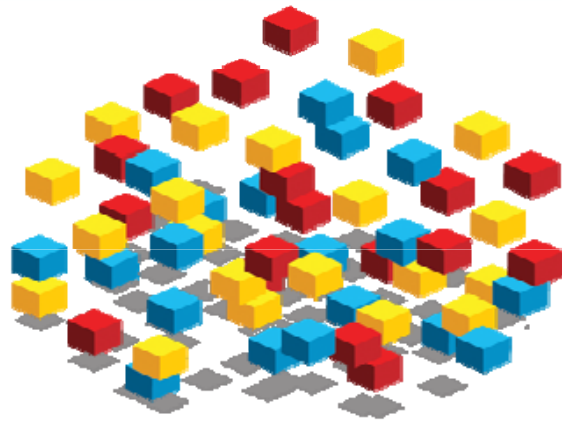
- Relates influencer to buyer
- Previous purchase history online for star wars lego
- Dave – bill to, Aiden ship to

Big Data

Discovers product likes and intent to purchase

The Relationship between MDM and Big Data

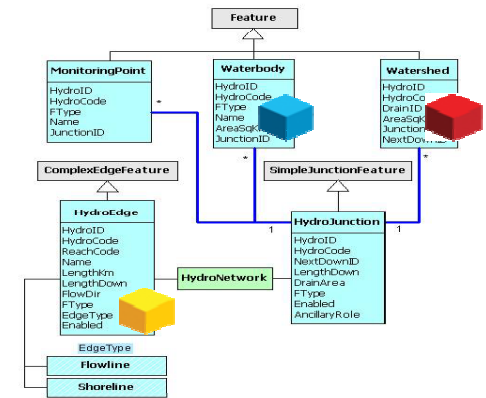
Big Data – Any Volume or Variety



Insights



MDM



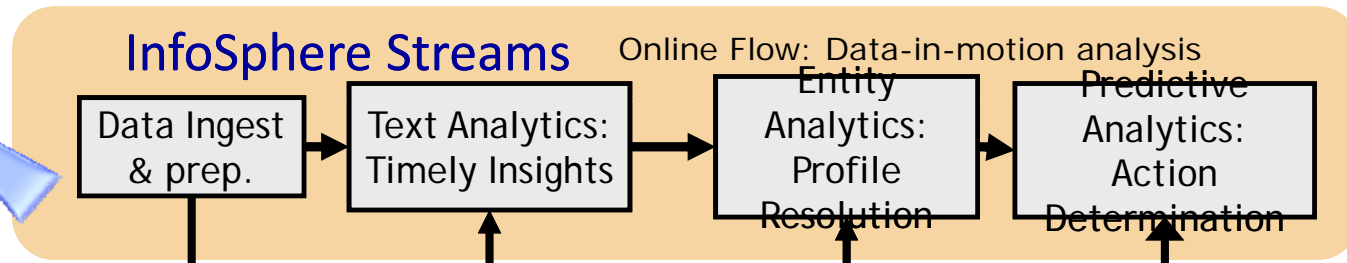
A deeper look at a typical solution architecture

Streaming content
external structured
and unstructured

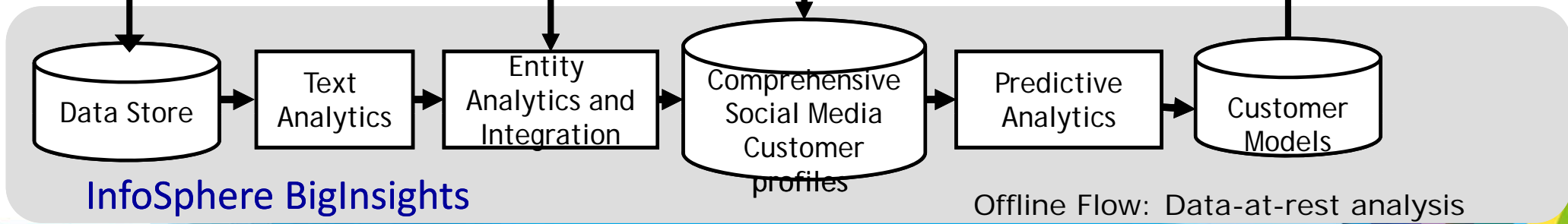
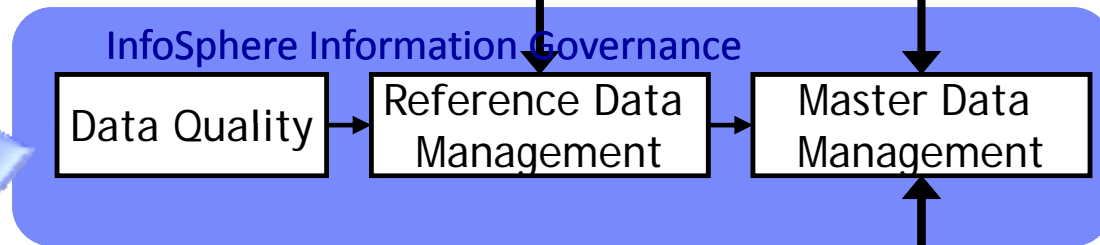


Big Data

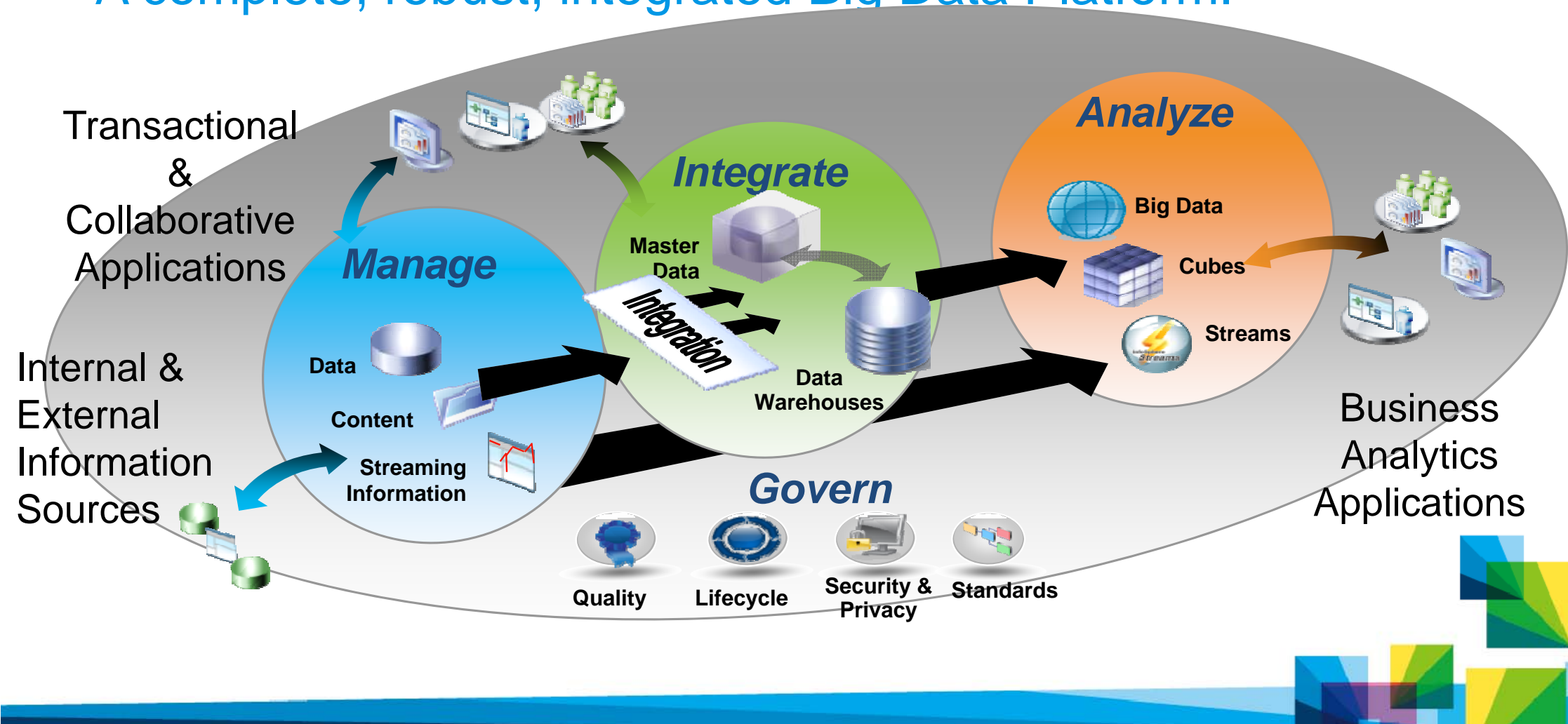
Internal
Structured data



Timely
Decisions



A complete, robust, integrated Big Data Platform.





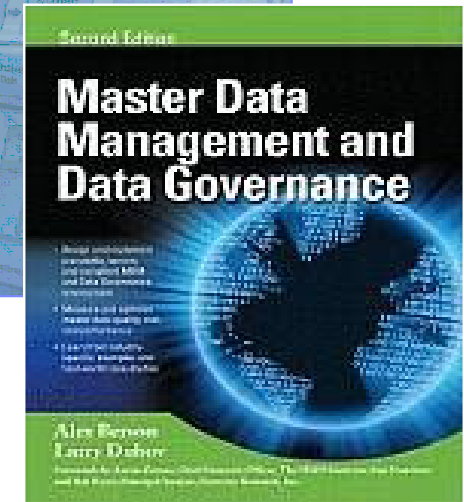
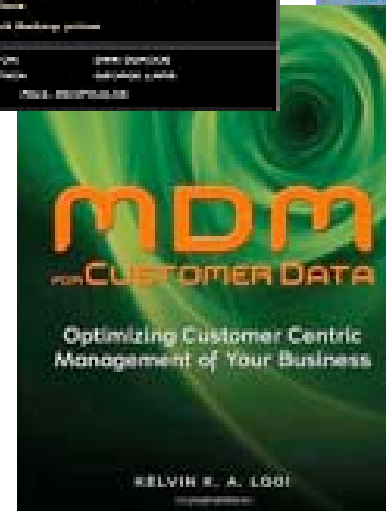
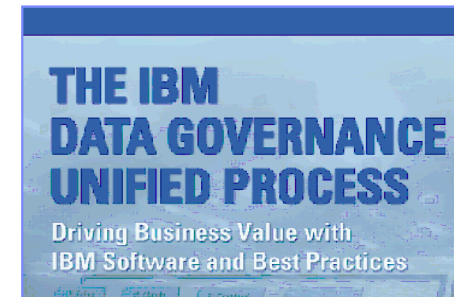
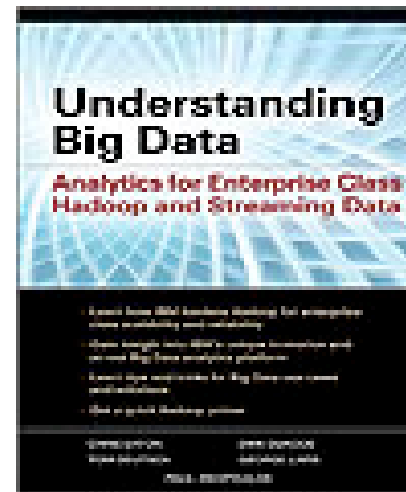
The bottom line

1. Find a business problem to solve
2. Enlist senior-level executives
3. Business needs to own it
4. It's ongoing, small value steps.
5. Define, measure, enhance and learn



How can your organisation achieve business insight with Big Data and MDM?

- Educate yourself on Big Data, Information Governance and MDM
- Talk to your IBM account manager or IBM business partner if you are interested in a Deep Dive workshop
- Talk to Tony, myself, the IBM and Partner team right now over coffee and beverages.





Big Data, Integration & Governance

28 - 30 August | Canberra; Melbourne; Sydney

Thank You