



Software
Summit 2010



Master Data Management & Mastering Unstructured Data

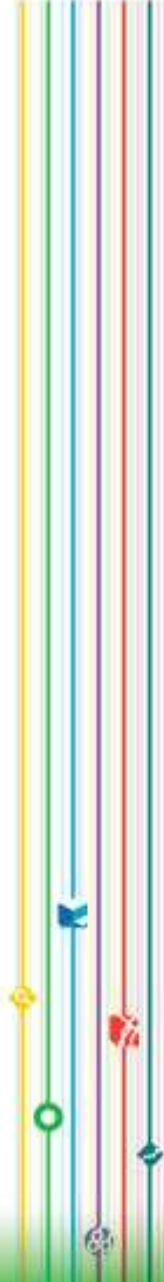


Agenda

- What is Master Data Management?
- What about unstructured data?
- A roadmap for Mastering Unstructured Data



MASTER DATA MANAGEMENT

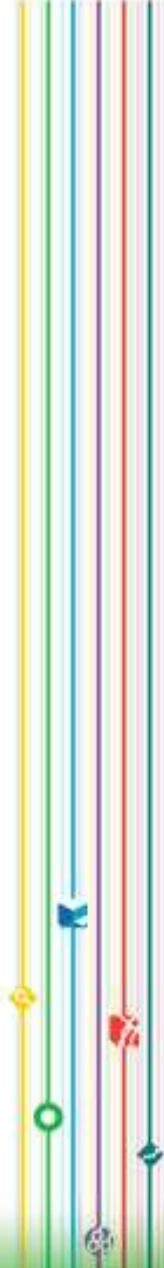


What is Master Data

- Master data is information that is key to the operation of business. This may include data about customers, products, employees, materials, suppliers, etc. which often turns out to be non-transactional in nature. Master data is often used by several functional groups and stored in different data systems across an organization and may or may not be referenced centrally; therefore, the possibility exists for duplicate and/or inaccurate master data.

What is MDM?

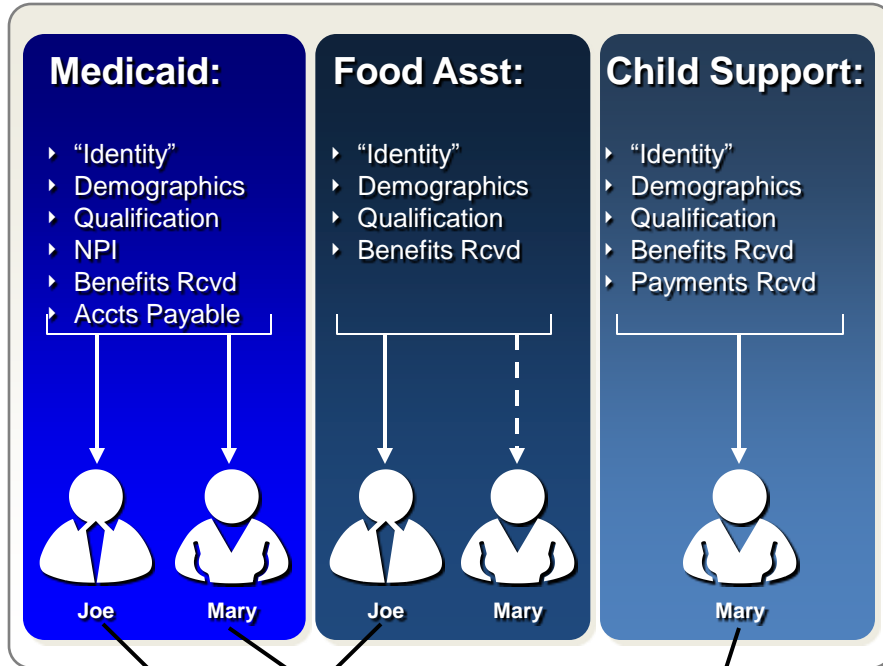
- Master Data Management (MDM) is a framework of processes and technologies aimed at creating and maintaining a “single version of truth”, an accepted system of record used across a diverse set of application systems, lines of business, and user communities.



Single Version of the Truth

Current State:

Account-Centric

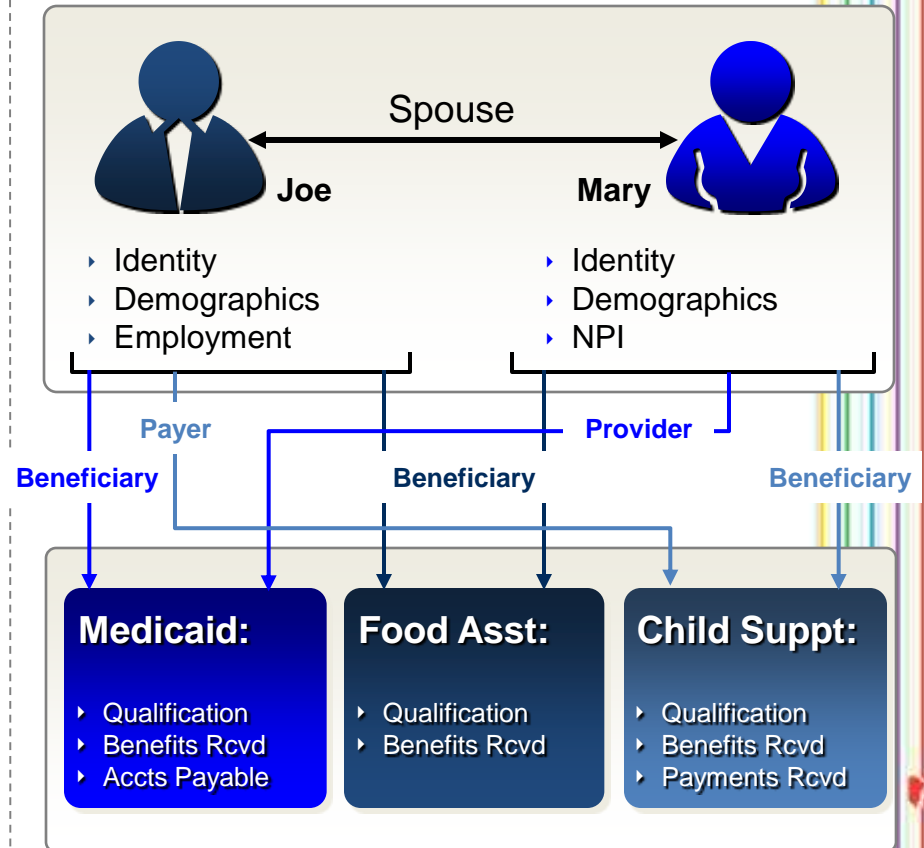


Derived



Future State:

Citizen-Centric



Program Grouping

We are at Work When You...

... Purchase Software



intuit.

... Pick Up Your Prescription



Walgreens

CVS/pharmacy
Expect something extra™

... Apply for a Loan



WELLS FARGO

Capital One

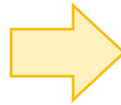
... Check In & Earn Points



IHG
InterContinental Hotels Group

CHOICE HOTELS INTERNATIONAL

... Apply for Health Insurance



UnitedHealthcare
Healing health care. Together.®

WELLPOINT

... File Insurance Claims



Allstate.
You're in good hands.

Grange
INSURANCE
Your partner in protection

... Register a Patient

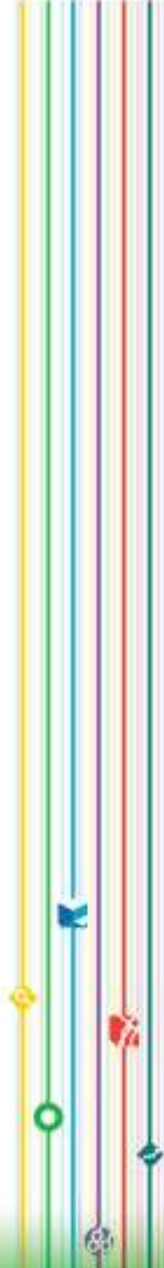


Sutter Health
With You. For Life.

BAYLOR
Health Care System

MDM Concepts

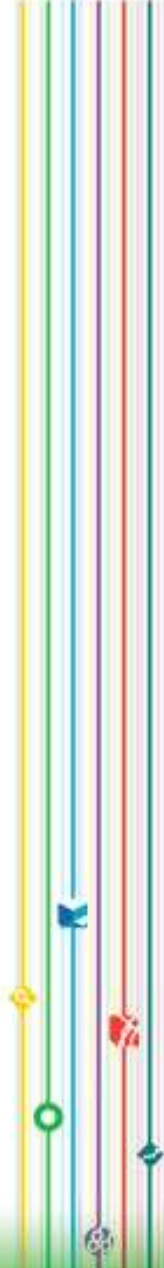
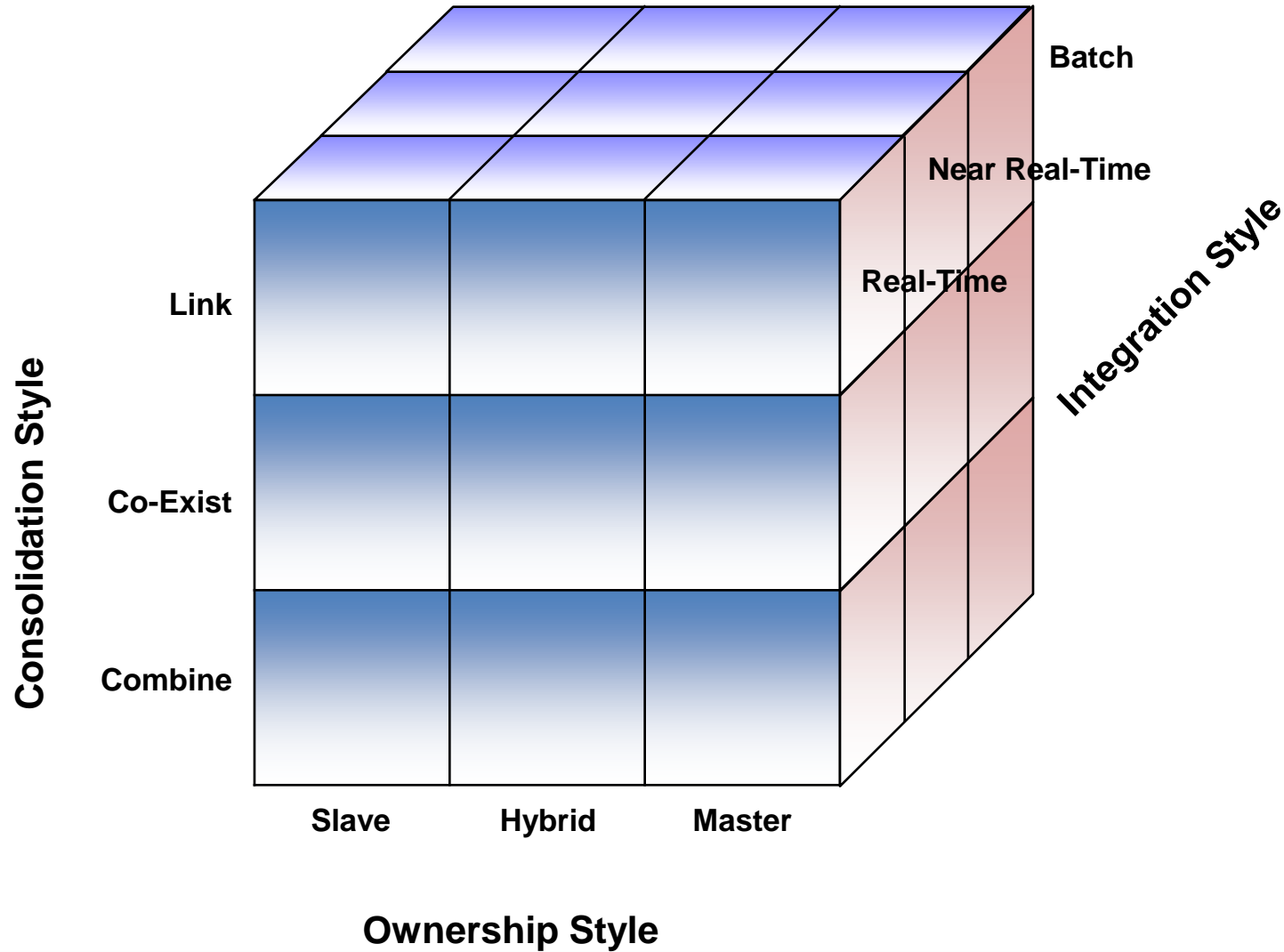
- Data modeling
- Integration & synchronization
- Security & audit
- Data quality



MDM Data Models

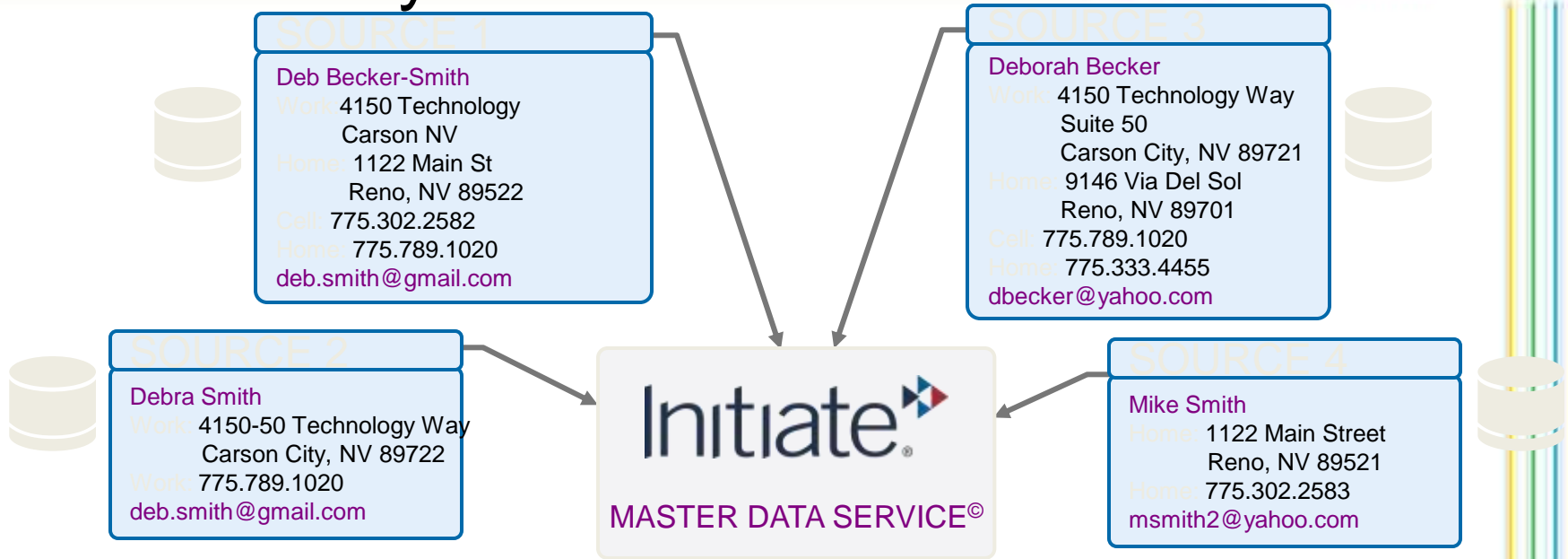
- **Approach 1: The “right” Data Model**
 - **Strategy:** To build a superset of data structures expecting that some data model fragments will work for the customer
 - Business Services
- ▶ **Approach 2: Metadata Model**
 - ▶ **Strategy:** To build a metadata layer that enables support for any data model
 - ▶ CRUD Services
- ▶ **Approach 3: Abstract Data Model**
 - ▶ **Strategy:** To build a model that supports any data modeling structures by translation DDL into DML within an abstract data model
 - ▶ Abstract CRUD Services

Integration & Synchronization



- Security controls:
 - Which sources a user can see
 - Which attributes a user can see
 - Which interactions a user can perform
- Audit
 - Who changed what data and when
 - Who saw what data and when
 - How and why records were matched

Data Quality



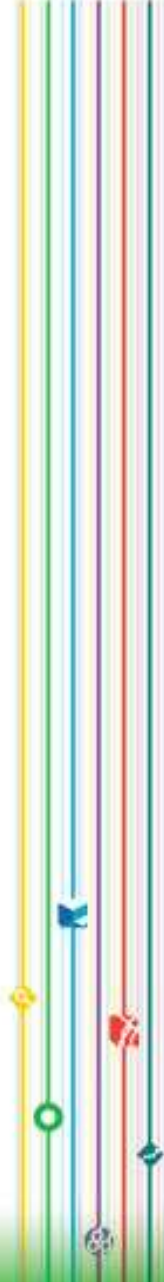
➔

COMPOSITE VIEW	Deborah Becker	1122 Main Street Reno NV, 89522	4150 Technology Way Suite 50 Carson City, NV 89721	775.333.4455	775.789.1020	775.302.2582	deb.smith@gmail.com
SOURCE ID	NAME	HOME ADDRESS	WORK ADDRESS	HOME PHONE	WORK PHONE	CELL PHONE	EMAIL
IDENT 23211	Debra Smith		4150-50 Tecnology Way Carson City, NV 89722		775.789.1020		deb.smith@gmail.com
DOS 32323	Deborah Becker	9146 Via Del Sol Reno, NV 89701	4150 Technology Way Suite 50 Carson City, NV 89721	775.333.4455		775.789.1020	deb.smith@gmail.com
ICE 10021	Deb Becker-Smith	1122 Main Street Reno NV, 89522	4150 Technology Carson NV	775.789.1020		775.302.2852	deb.smith@gmail.com

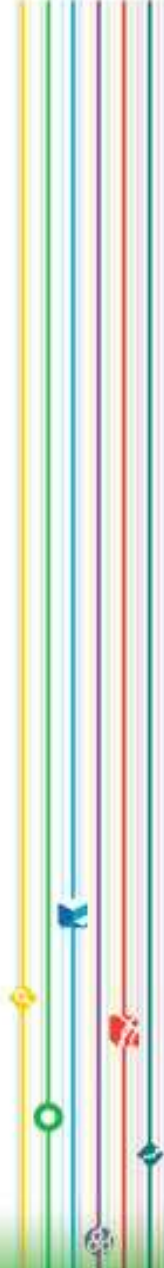
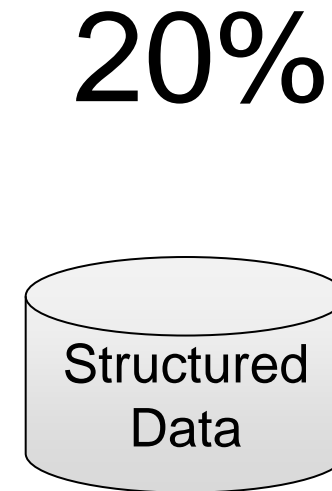
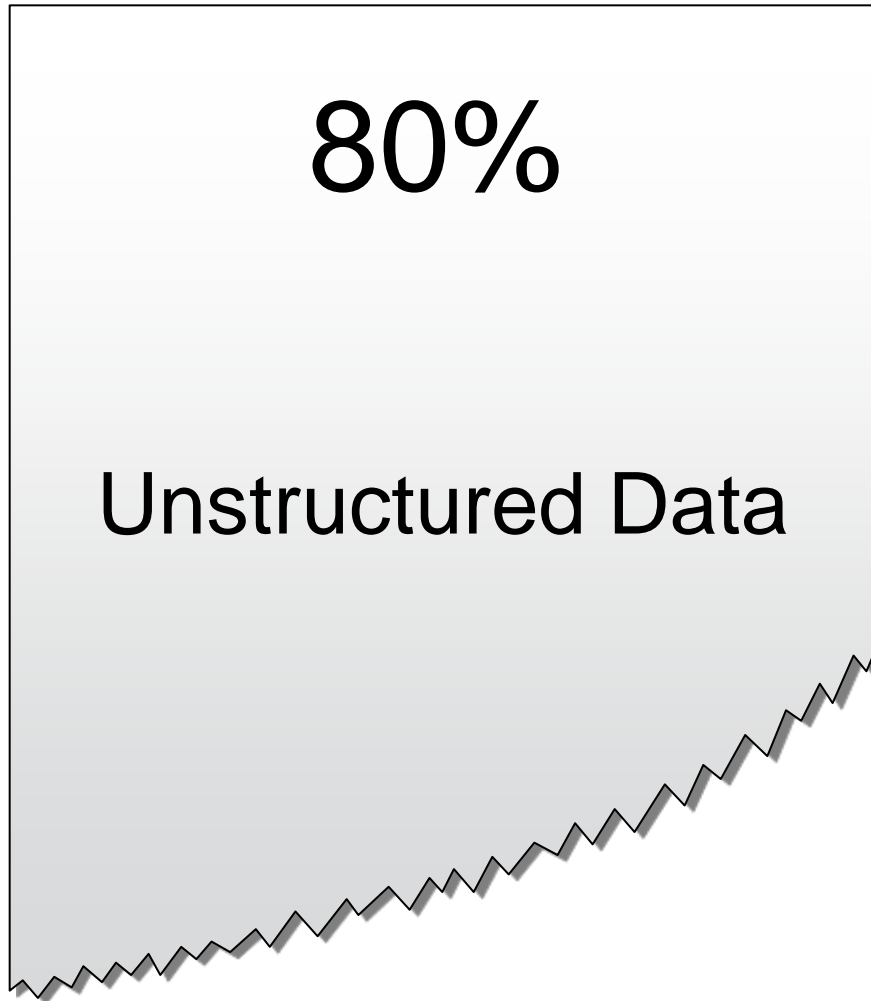
MDM DEMO



UNSTRUCTURED DATA



Unstructured Data

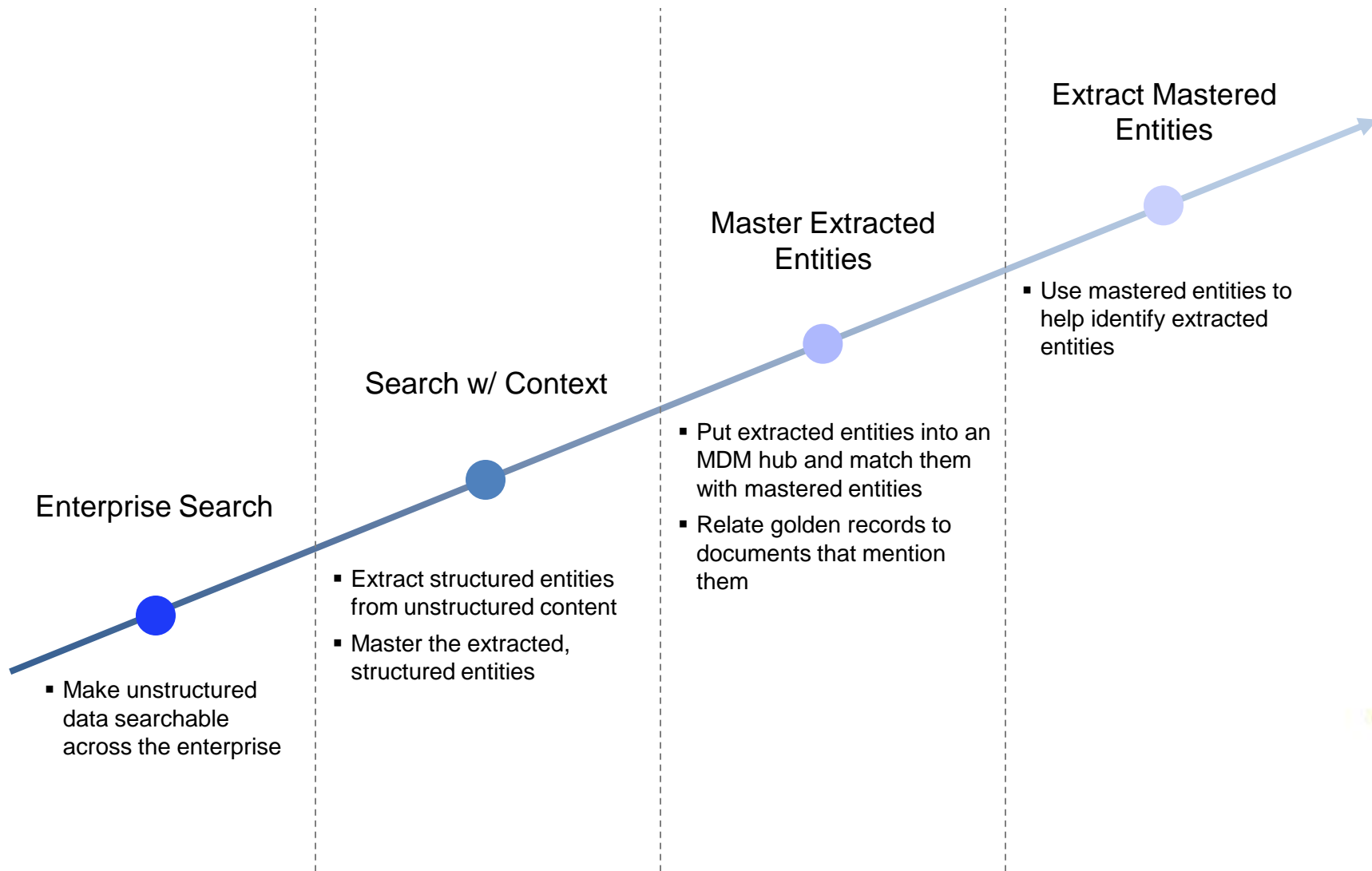


Example

At about 1040 hours on **5th April 2001**, I met with **Ms. Vanessa Price** at **61 South Chorley Drive** regarding a vehicle theft. Ms. Price said she parked her car by a parking meter outside Chorley Leisure Centre at about 0945 hours and went into a nearby shop to return a faulty torch she had purchased the previous day. She said that when she returned to the leisure centre at about 1000 hours, she discovered her car was missing.

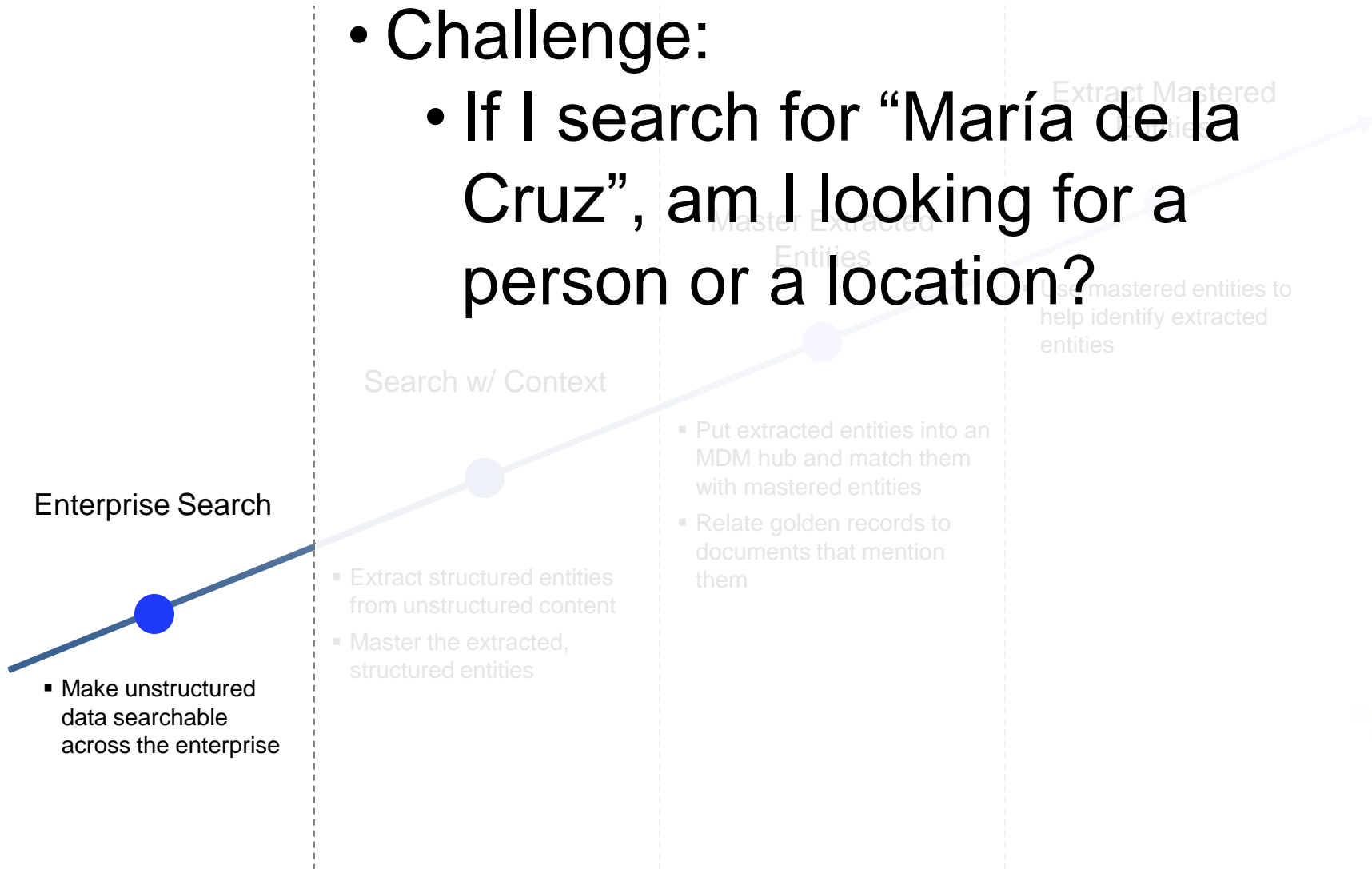
Ms. Price described her car as a maroon, **1999 Mitsubishi Eclipse** with a black convertible roof. The car registration number is **GTL-682-P**. She estimated the value of the car at £8,500 and said there were no distinguishing marks or items.

Roadmap

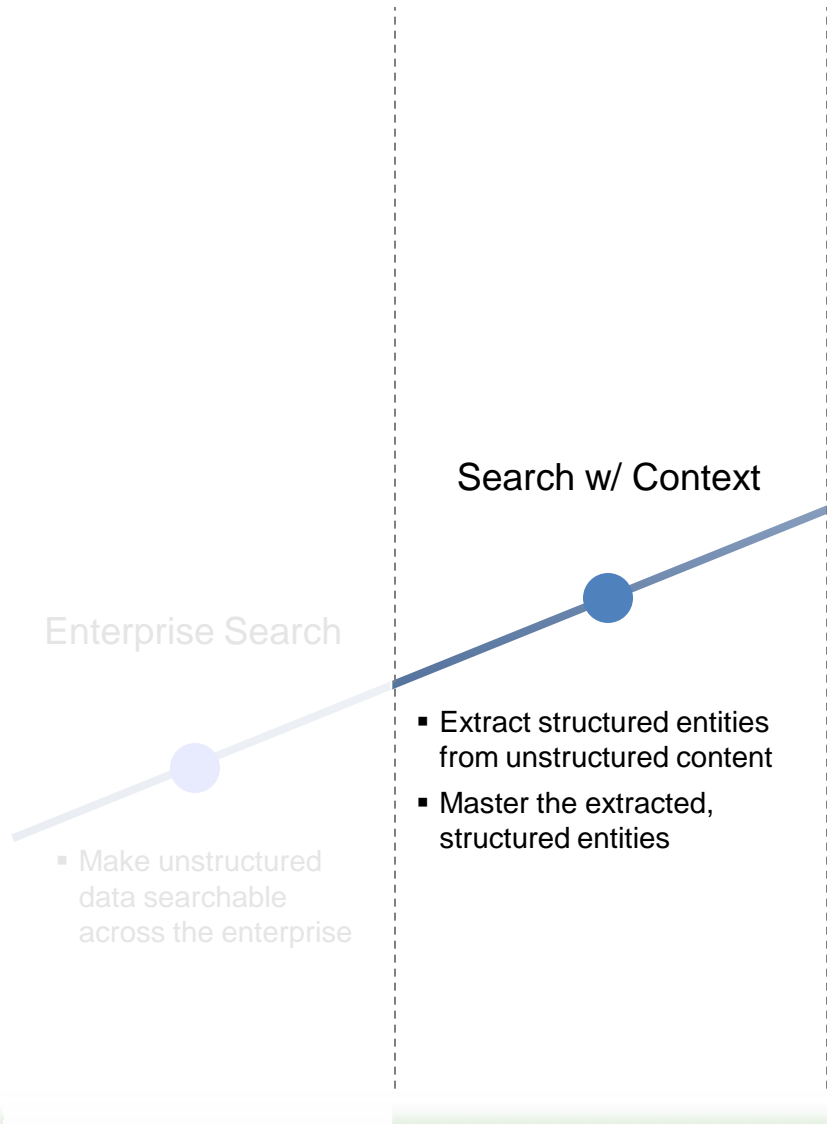


1st Generation: Enterprise Search

- Challenge:
 - If I search for “María de la Cruz”, am I looking for a person or a location?

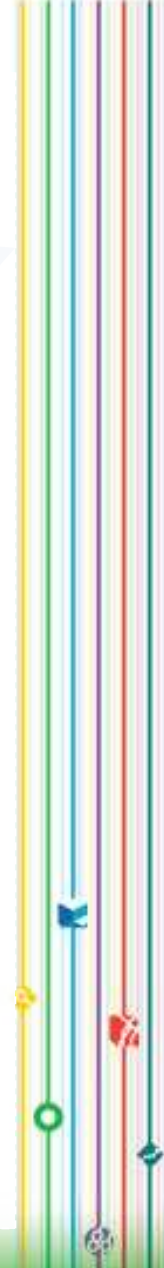


2nd Generation: Search w/ Context



- Challenge:
 - How many unique individuals are represented in the search results

- Put extracted entities into an MDM hub and link them with mastered entities
- Relate golden records to documents that mention them



How Many Nick Denning's Are There?

Nick Denning

Managing Director at Diegesis Limited
Information Technology and Services |
London, United Kingdom

Managing Director at Diegesis Limited, Chief
Technical Officer and before that Chairman
and before that Chairman and Managing
Director at Strategic Thought Group Plc

Nick Denning

Assistant Manager at the Federal Reserve
Bank of Kansas City
Banking | Kansas City, Missouri Area

Assistant Manager at Federal Reserve Bank
of Kansas City, University of Missouri-Kansas
City, Central Missouri State University

Nick Denning

Owner, Procol Ltd and Facilities Services
Consultant
Facilities Services | Swindon, United Kingdom

Nick Denning

Officer at HM Forces
Military | Colchester, United Kingdom

Nicholas Denning

National Commercial Manager at Aggregate
Industries

Construction | Bath, United Kingdom

National Commercial Manager at Aggregate
Industries, Submissions Manager at CEMEX

Nick Denning

Independent Entertainment Professional
Entertainment | Greater Los Angeles Area

Nick Denning

Financial Services | Greater Chicago Area

Nick Denning

Finance Director at Precision Colour Printing Ltd
Printing | Telford, United Kingdom
Finance Director at Precision Colour Printing Ltd

How Many Nick Denning's Are There?

Electoral Roll 2008-09 Nicholas I Denning Age guide: 50-54, Newport, Gwent, NP11

Electoral Roll 2006-09: Nicholas A Denning Age guide: 30-34, Bath, Avon, BA1

Electoral Roll 2005 Director Report: Nicholas Hugh Denning Age: 56, Chippenham, Wiltshire, SN15

Electoral Roll 2004-06: Nicholas C Denning, Stafford, Staffordshire, ST21

Electoral Roll 2003-05: Nicholas J Denning, Lowestoft, Suffolk, NR32

Electoral Roll 2007-09: Nicholas R Denning Age guide: 17-20, Bristol, Avon, BS7

Electoral Roll 2002: Nicholas C Denning, Rugeley, Staffordshire, WS15

Electoral Roll 2003-06: Nicholas P Denning, Guildford, Surrey, GU2

Electoral Roll 2002: Nicholas Denning, Dunmow, Essex, CM6

Electoral Roll 2002: Nicholas J Denning, Bristol, Avon, BS3

Electoral Roll 2006-09: Nicholas R Denning Age guide: 40-44, Drybrook, Gloucestershire, GL17

Electoral Roll 2004: Nicholas H Denning, Chippenham, Wiltshire, SN14

Electoral Roll 2002-03: Nicholas H Denning, Chippenham, Wiltshire, SN14

Electoral Roll 2005: Nicholas Denning, Bath, Avon, BA1

Director Report: Nicholas John Denning Age: 36, Bristol, Avon, BS5

Electoral Roll 2002: Nicholas H Denning, London, SW19

Electoral Roll 2002: Nicholas Denning, Dunmow, Essex, CM6

Electoral Roll 2006: Nicholas Denning, Cheltenham, Gloucestershire, GL50

Electoral Roll 2009: Nick T Denning Age guide: 40-44, Newport, Isle Of Wight, PO30

Electoral Roll 2004: Nicholas Denning, Bath, Avon, BA1

Electoral Roll 2003-06: Nicholas G Denning, Colchester, Essex, CO3

Electoral Roll 2004-05: Nick Denning, London, SE16

Electoral Roll 2008: Nicholas Denning, Lowestoft, Suffolk, NR33

Electoral Roll 2002: Nicholas A Denning, Bristol, Avon, BS16

Electoral Roll 2002-09: Nicholas Denning Age guide: 35-39, Exeter, Devon, EX4

Electoral Roll 2002-09: Nicholas J Denning Age guide: 45-49, Kingswinford, West Midlands, DY6

Electoral Roll 2002-05 Director Report: Nicholas Richard Denning Age: 44, Drybrook, Gloucestershire, GL17

Electoral Roll 2002-07: Nicholas I Denning Age guide: 46-50, Newport, Gwent, NP11

Electoral Roll 2002-06: Nicholas A Denning, Ascot, Berkshire, SL5

Electoral Roll 2007-09 Director Report: Nicholas Charles Denning Age: 46, Stafford, Staffordshire, ST21

Director Report Nicholas Henry Denning Age: 52, London, SW19

Source: 192.com

3rd Generation: Master Extracted Entities

- Challenge:
 - Entity extraction is imperfect
 - Which name goes with which birth date?

“Alex Eastman and Matt Eastman are brothers. The latter was born on 2 Feb 1969. The former was born on 14 Sep 1973.”

▪ Extract structured entities from unstructured content

Master Extracted Entities

- Put extracted entities into an MDM hub and match them with mastered entities
- Relate golden records to documents that mention them

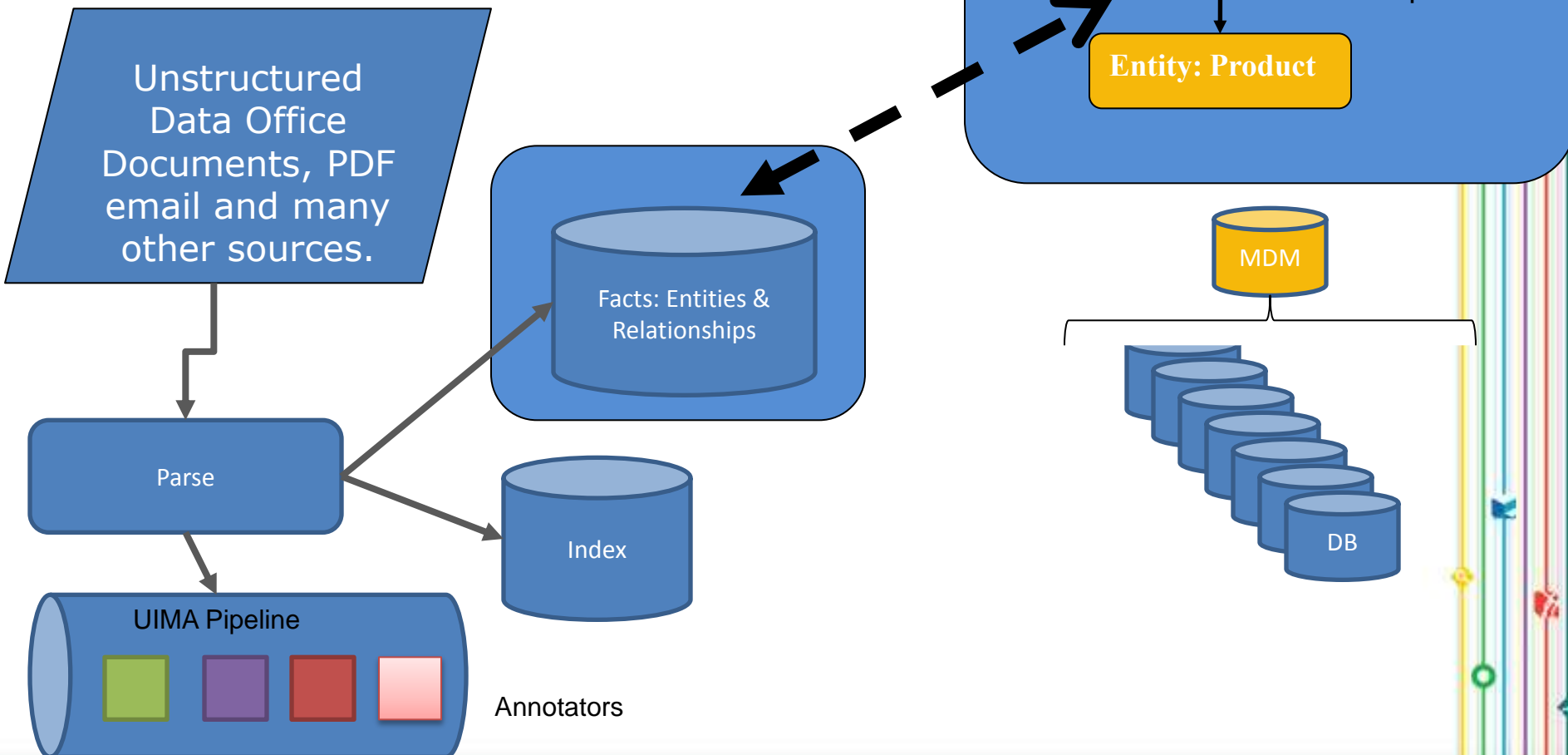
Extract Mastered Entities

- Use mastered entities to help identify extracted entities



Entity Extraction

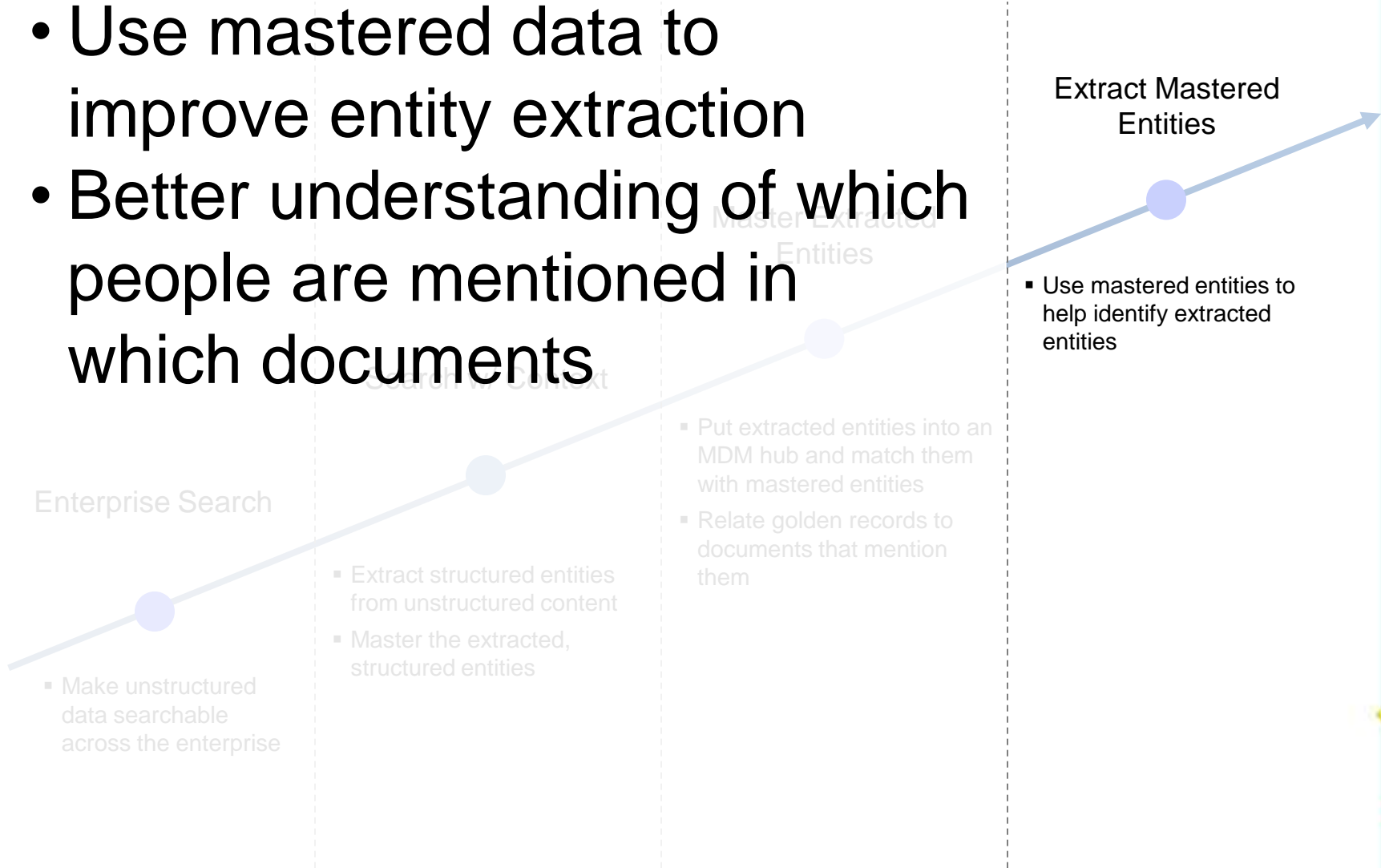
How to link entities in unstructured information to MDM gold records?



4th Generation:

Extract Mastered Entities

- Use mastered data to improve entity extraction
- Better understanding of which people are mentioned in which documents



Extract Mastered Entities

43, Lingfield Road, SW19 4PZ	Nicholas Henry Denning	07710338072	02089471722
45, Elmbourne Road, SW17 8JS	Nick Denning	07710338072	02089471722
37, Evelyn Road, SW19 8NT	Nicholas Denning		02082867587
4, Welford Place, SW19 5AJ	Nicholas Henry Denning		02082866578
43, Lingfield Road, SW19 4PZ	N H Denning		01454412199
			02085436644

To: Nick Denning
From: Ben Yefferson

Tel: 02085436644
Addr: 37, Evelyn Road
SW19 8NT
Addr: 15 Garden Crescent
London N1 WATT

Entity extraction

- Address: 37, Evelyn Road, SW19 8NT
- Address: 15, Garden Crescent, N1 WATT
- Name: Nick Denning
- Name: Ben Yefferson
- Telephone: 02085436644

IBM

Software
Summit 2010 

Thank You

