AI and Master Data Management: A Synergy for Success

boomi

Dr. Madassar Manzoor

Solution Architect
Onepoint
https://www.onepointltd.com/

Tom Clarke

Data Management Solutions Consultant
Boomi

A BIG QUESTION

How can my company avoid the risks of implementing AI with reckless abandon?

Agenda

01 Rapidly evolving AI applications landscape.

02 Negative impact of poor master data.

03 Best practice for creating and maintaining accurate master data records.

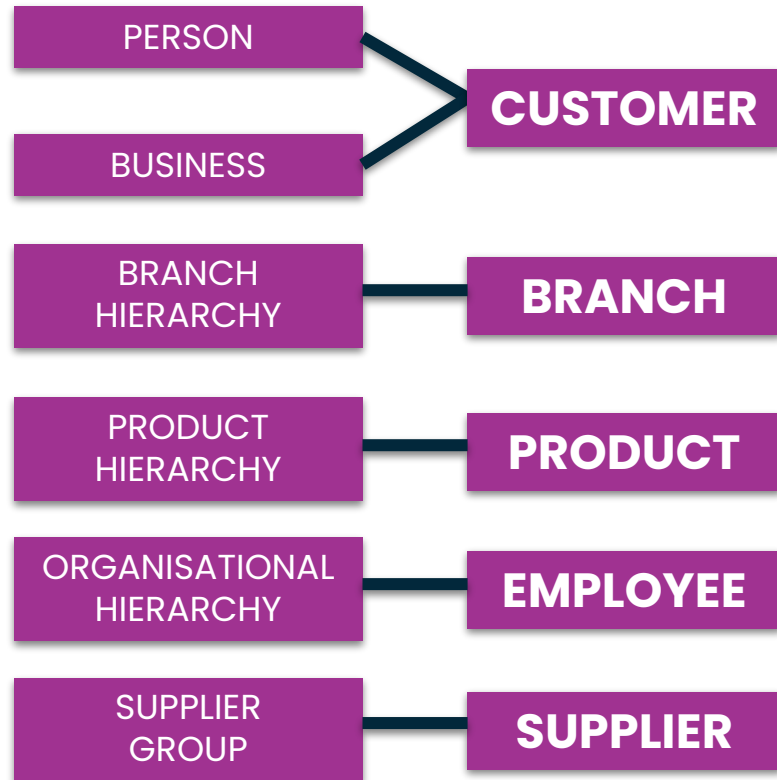**04** Use of AI to accelerate cleaning and matching of master data records.

05 Overview of Boomi's AI & Master Data Management capabilities.

Headlong charge to embrace AI

1	**CUSTOMER SERVICE** - Natural Language chatbots → any query, any time, no hand-offs
2	**DOCUMENT PROCESSING** – Application Processing, Supporting Documents, Contracts
3	**PREDICTIVE MODELS** – Sales, Marketing, Risk, Fraud, AML, Demand Forecasting …
4	**MASS PERSONALISATION** - Telematics, Wearable Devices, Avatars
5	**HUMAN RESOURCES** - Recruitment, On-boarding, Reviews, Development Plans
6	**FINANCE** - Accruals, Allocations, Reconciliation - Order-Delivery-Invoice Matching
7	**SENSOR, IMAGE & VIDEO DATA** - Predictive Maintenance, Quality Control, On-Shelf Availability …

BUT ... building AI applications on weak foundations

PERSON	
BUSINESS	**CUSTOMER**
BRANCH HIERARCHY	**BRANCH**
PRODUCT HIERARCHY	**PRODUCT**
ORGANISATIONAL HIERARCHY	**EMPLOYEE**
SUPPLIER GROUP	**SUPPLIER**

Universal Weaknesses in Master Data

1. **RECORDED** in multiple systems or spreadsheets
2. Held in **DIFFERENT** table and file structures
3. Using **DIFFERENT** reference data codes
4. NO controls to prevent **DUPLICATES**
5. NO controls to **VALIDATE** data
6. NO systematic way to record **HIERARCHIES**
7. NO **SYNCHRONISATION** across systems
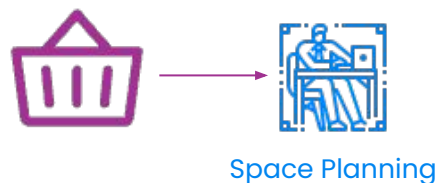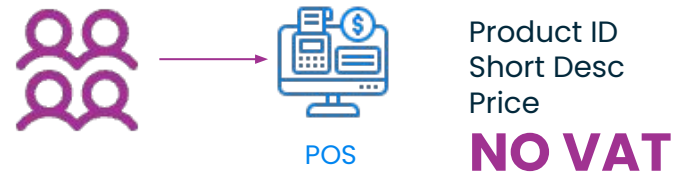
The impact of poor master data: case study

Names on List	96
Duplicates	20
No mortgage	13
Redeemed mortgage	19
No telephone number	4
Invalid telephone number	7
New sales – in branch	13
New sales – via partner	1
New sales – via competitor	4
Unable to contact	5
SALES	7

Client **DID NOT** Accept the need to create a **SINGLE VIEW OF CUSTOMER**

7% Take-Up Rate
Considered **'great'** success

The impact of poor master data: case study

Perception versus reality

Product ID
Style, Size, Season,
Colour, Fabric,
Order Quantity,
Delivery Date

Product ID
Short Desc
Price
NO VAT

POS

Product ID
Store ID
Hanging
Packaged

Space Planning

Product ID
Hanging
Package Size
Pre-Processing

Fashion Retail
EUR 10 billion, 14 Countries

VAT Receipt
request → Email to Finance

VAT Return
Calculation →

Financial System
Country, Total Sales

Enterprise Data Warehouse
Country, Department, Class, Total Sales

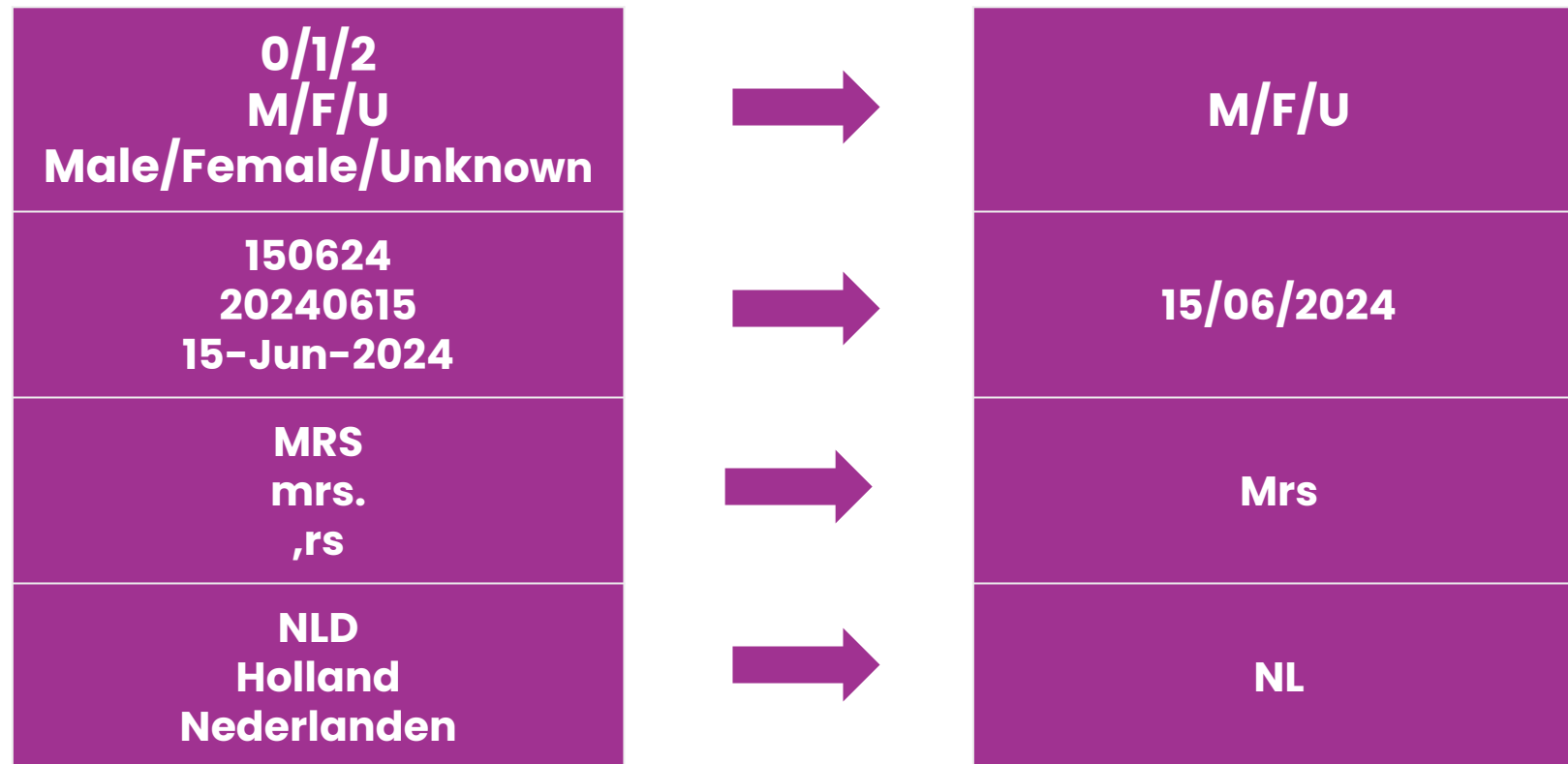8

Challenge #1: 'Dirty' data is everywhere

Found in EVERY table on EVERY system

Health Service Provider

GBP 250 million pa
160k Customer Records

Records	Contents of 'title' field (free text field)
755	DISTINCT VALUES
53,614	BLANK (NULL) □ 33.8% of records
	..... or ---- or +++
	,Oss or ,R or ,Rs etc
	Refugee, Social Worker, Duty Nurse, Head Teacher etc
	Carer, Wife, Husband, Stepfather, Sister, Cousin
	Anne, Barbara, Craig, Dylan etc
	21/10/1997 etc
	199 High Street etc, Lives with Daughter

Challenge #2: Reference data differs across system

Reference data will need to be translated into a common form = STANDARDISED

0/1/2 M/F/U Male/Female/Unknown	M/F/U
150624 20240615 15-Jun-2024	15/06/2024
MRS mrs. ,rs	Mrs
NLD Holland Nederlanden	NL

Challenge #3: Duplicate and inconsistent records

Dictates the need for SOPHISTICATED capabilities to MATCH & MERGE records

SYSTEM	TITLE	FIRST NAME	FAMILY NAME	DATE OF BIRTH	GENDER	EMAIL
A	Mr	Madassar	Manzoor	01-Jan-97	M	mm@gmail.com
A	Dr	Madassr	Maznoor	10-01-1997		mm@gmail.com
A		Madasser	Manzor	1979/01/01		07711-223344
B	M.R.	NANSOOR	Madassar	01/10/97	mail	nn@gnail.com
C	Manzoor, Madassar/MR/Employed/HomeOwner/19970101/MALE/mm@gmail.com					

Master Data Management – Best Practice

The **DESIGN** process

SOURCE SYSTEMS

EXTRACT

DATA PROFILING
Every field in every file

DEFINE CLEANSING RULES

DEFINE STANDARDISATION RULES

ASSES ENRICHMENT REQUIREMENTS

What is data enrichment?

Property Occupancy

Royal Mail

Companies House

CHARITY COMMISSION
FOR ENGLAND AND WALES

Yell™

Master Data Management Sytem

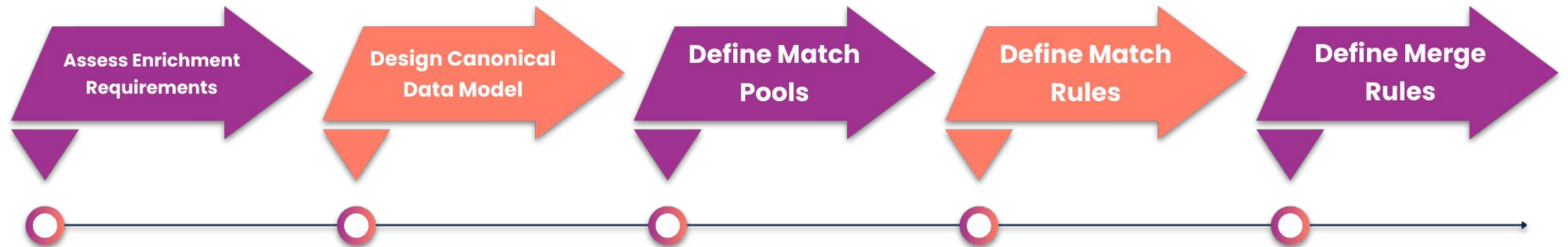| Ingest | Clean **Enrich** Match Merge | Golden Records |

Financial Analysts

IDENTIFIED:
- **1600** businesses in 'empty' business premises
- **2700** businesses in residential properties
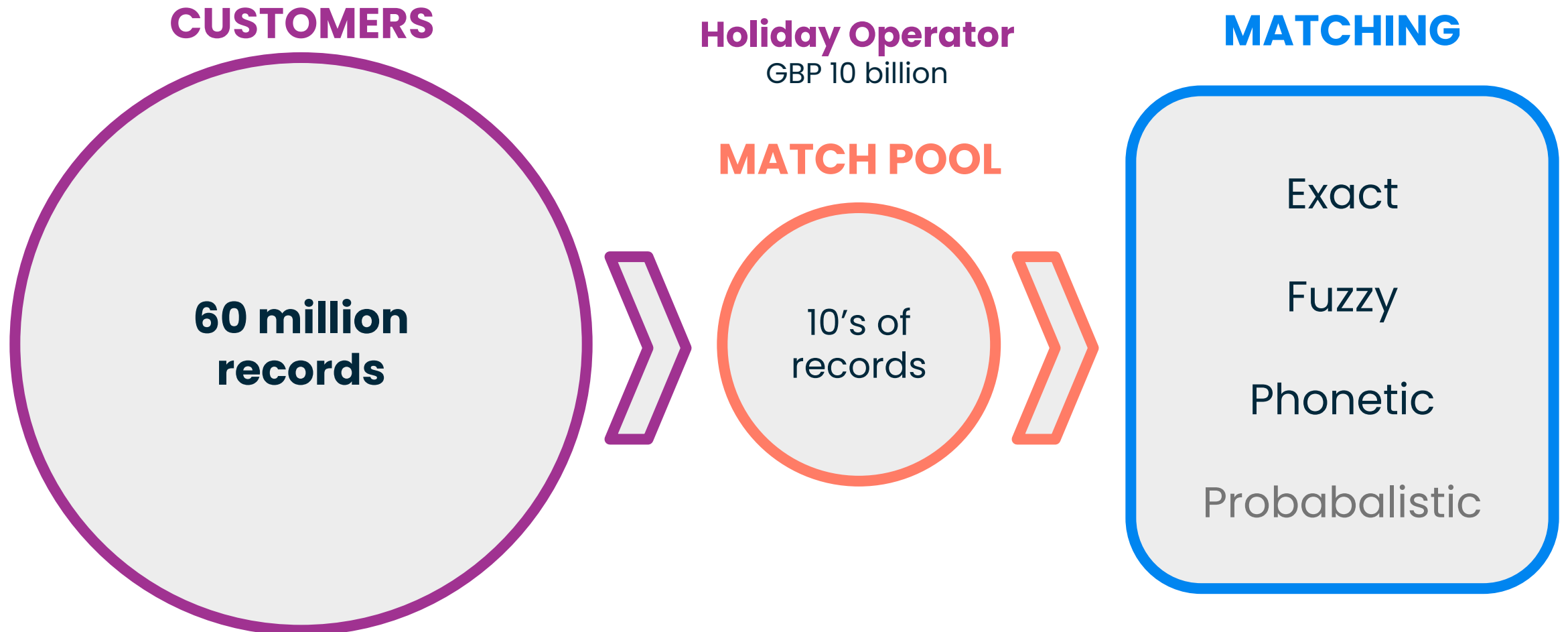- Potential to increase income by over **25%**

13

Master Data Management – Best Practice
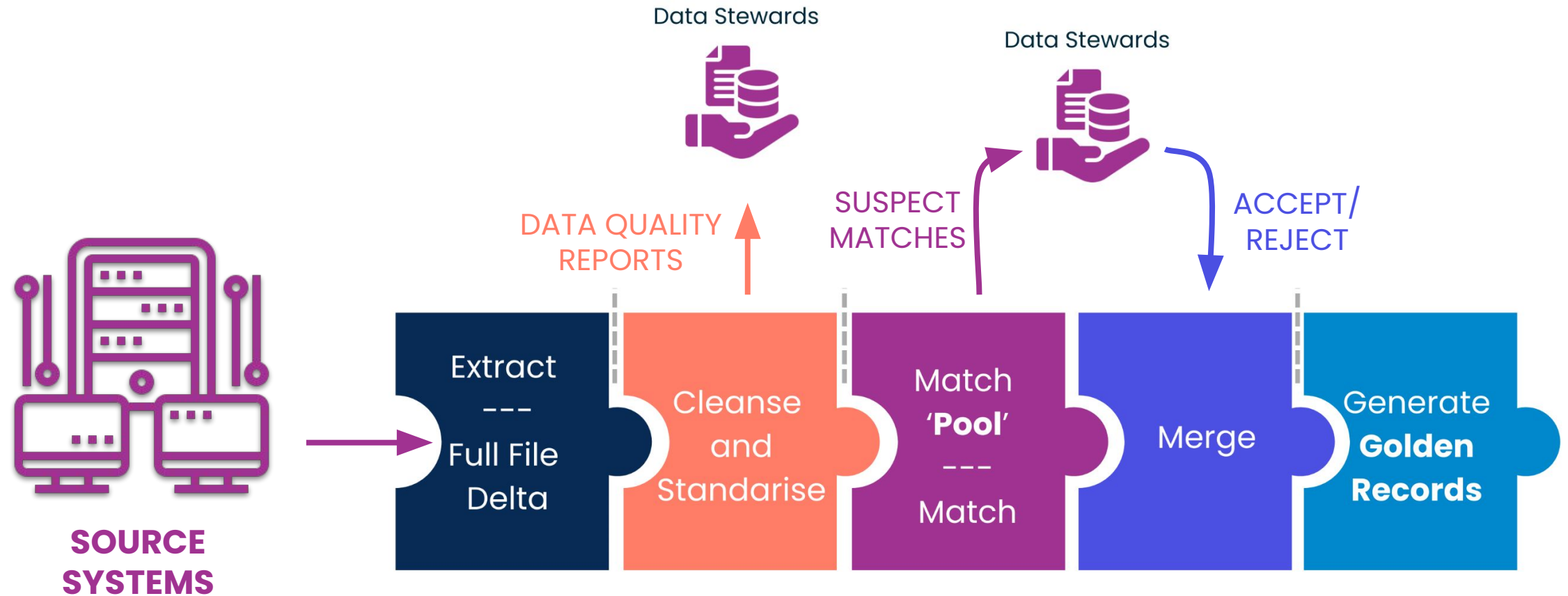
The **DESIGN** process … continued

Assess Enrichment Requirements → Design Canonical Data Model → Define Match Pools → Define Match Rules → Define Merge Rules

Master Data Management – Best Practice

AUTOMATICALLY generate match pools, **AUTOMATICALLY** match records

CUSTOMERS

Holiday Operator
GBP 10 billion

MATCHING

MATCH POOL

60 million records

10's of records

Exact

Fuzzy

Phonetic

Probabalistic

Master Data Management – Best Practice

The **DEPLOYMENT** process

Data Stewards

Data Stewards

DATA QUALITY
REPORTS

SUSPECT
MATCHES

ACCEPT/
REJECT

SOURCE
SYSTEMS

Extract

Full File
Delta

Cleanse
and
Standarise

Match
'**Pool**'

Match

Merge

Generate
**Golden
Records**

16

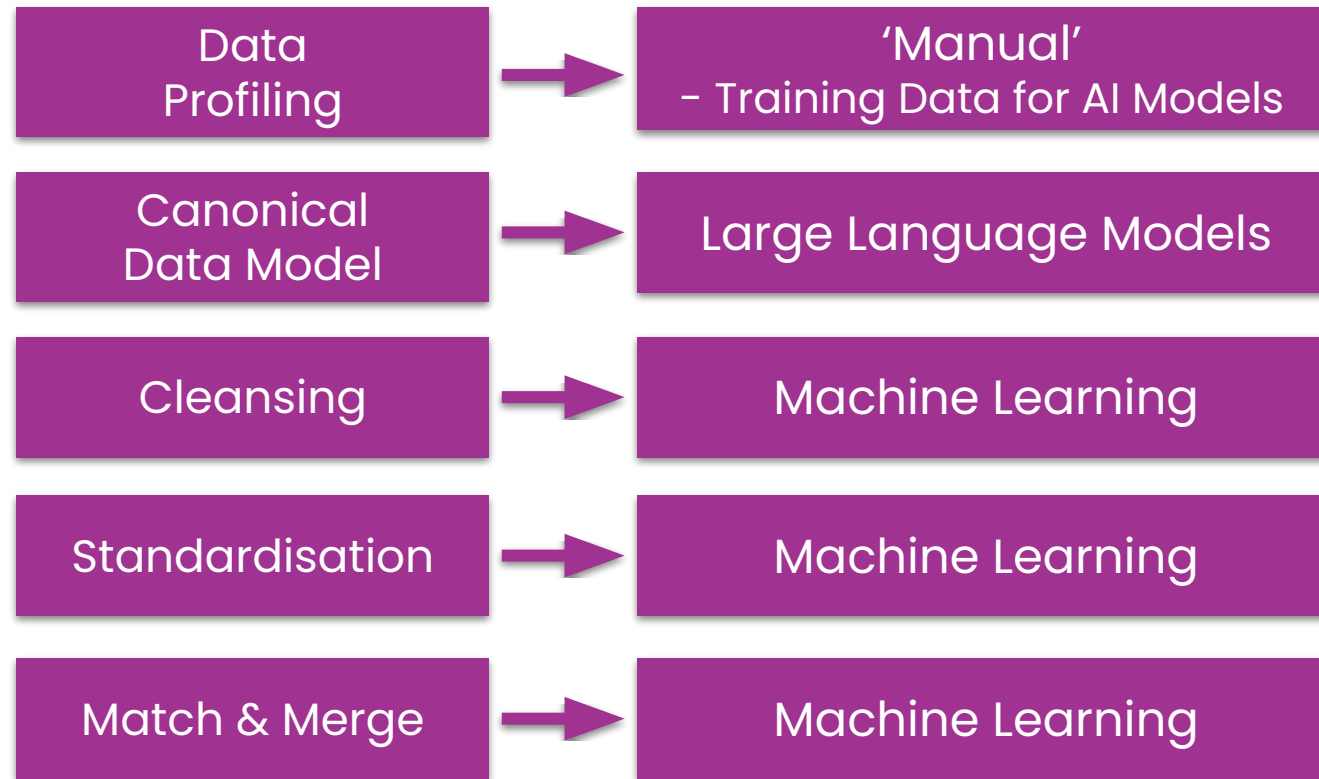# Master Data Management – Best Practice

Closing the loop - **SYNCHRONISE** master data across operational systems

Solar Energy Generator
160 Solar Farms

Boomi Integration

Ingest — Clean Match Merge — Synchronise

Site
People
Project
IT System
Legal Entity

Boomi Master Data Hub

Quarantined Records

Data Stewards

quickbase
Site and Asset (API)
Project (API)

Azure Active Directory
People (API)

SAP SuccessFactors
People (API)

monday.com
IT Systems (API)

OneTrust
PRIVACY, SECURITY & TRUST
IT Systems (API)

Legal Entity (API)
Shareholder (API)

melon
powering CORPORATE TRAVELLER

SAP Concur

Culture Amp

GIS
GEOGRAPHICAL INFORMATION SYSTEMS

IFS

Microsoft Azure
Blob Storage

Master Data Management – emerging capabilities

Using AI to ACCELERATE development and IMPROVE accuracy

Data Profiling	→	'Manual' – Training Data for AI Models
Canonical Data Model	→	Large Language Models
Cleansing	→	Machine Learning
Standardisation	→	Machine Learning
Match & Merge	→	Machine Learning

ANOTHER BIG QUESTION

How can my company profit from the AI Revolution, instead of being left behind?

Data is Everything

- Select an initiative based on cleanest or best data

- Garbage in / garbage out still applies with AI
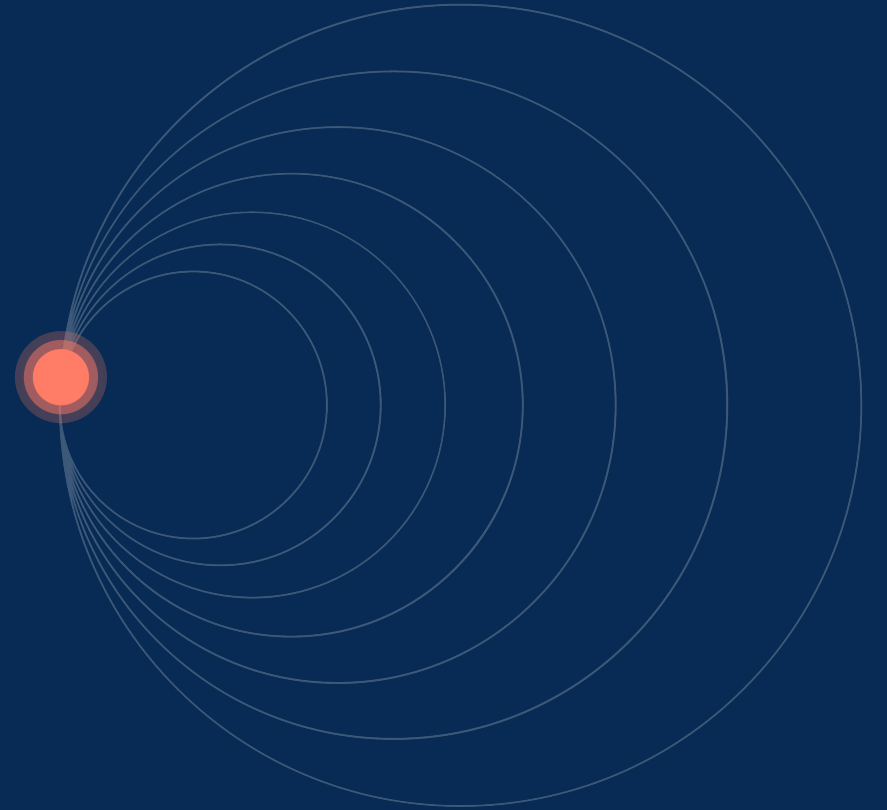
- 70% of effort is working on the data

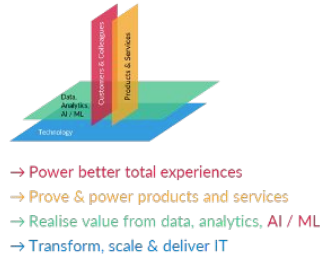# Boomi's Unified Platform
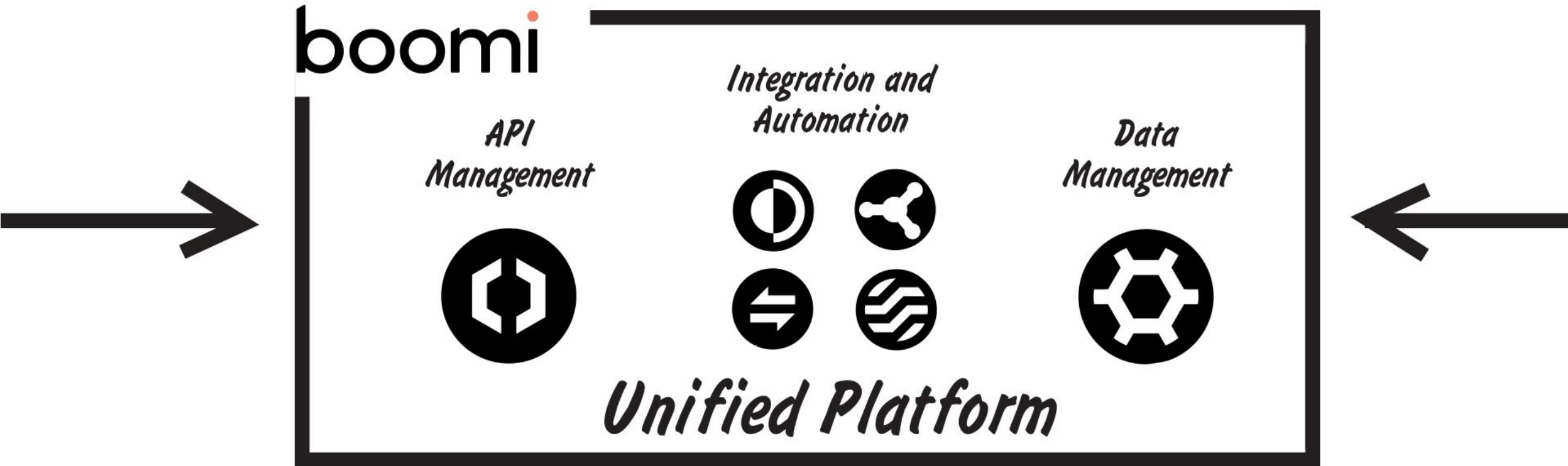
Boomi's Unified, AI Ready Platform

- 360+ Applications
- 1000+ Databases
- 10,000+ APIs

Interconnected Process

boomi

API Management · **Integration and Automation** · **Data Management**

Unified Platform

iPaaS

One Platform
Integration & Automation

AI

Models & Agents

Data Storage
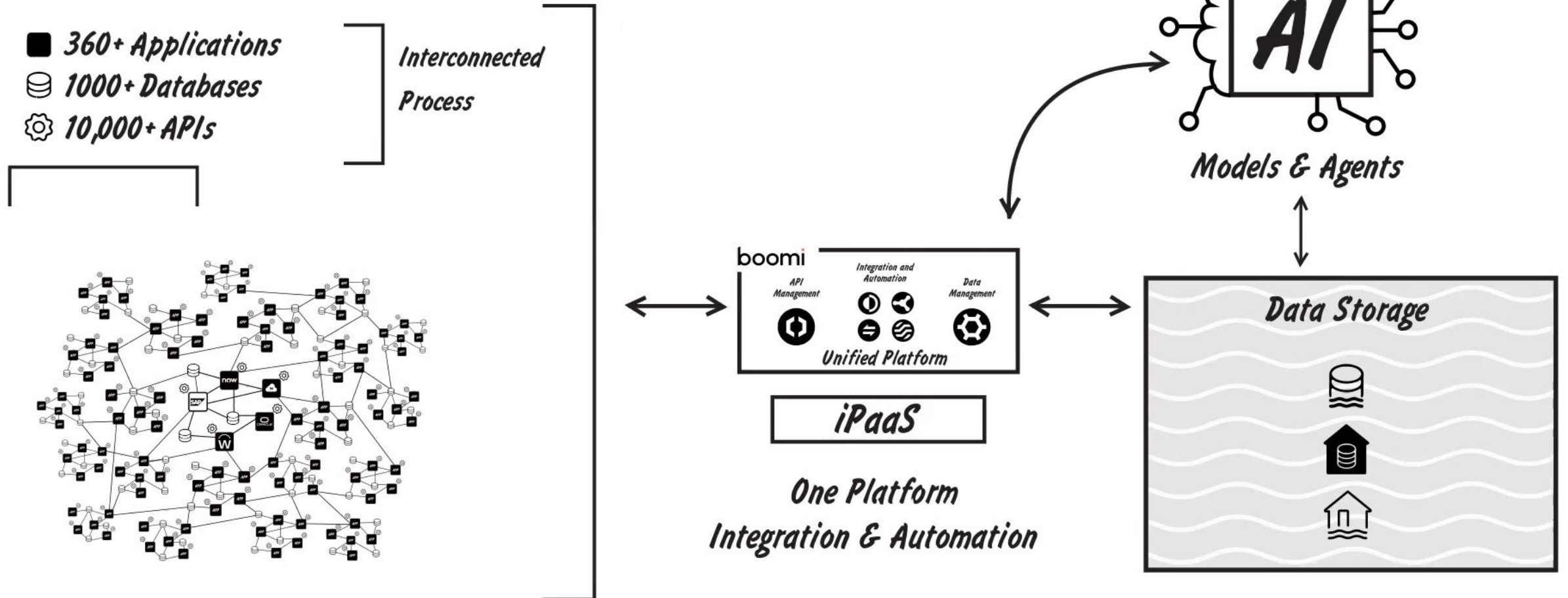
Thank you

We are Boomi.

Be Bold. Be You. Be Boomi.

Introducing Onepoint

https://www.onepointltd.com/

What we do

We **architect, prototype, build, and manage** data and AI powered solutions around four strategic digital priorities. We partner with clients looking for high-impact, enterprise-grade advice and IT services to realise their most critical digital transformations.

→ Power better total experiences
→ Prove & power products and services
→ Realise value from data, analytics, AI / ML
→ Transform, scale & deliver IT

Who we serve

Our primary buyers are **business, technology, and data / analytics leaders** looking for deep technical experts who 'speak business' and know how to 'get to business value' quickly. Our expertise is transferrable and relevant to **public and private sectors.**

FLIGHT CENTRE

Lightsource bp

MEGGiTT
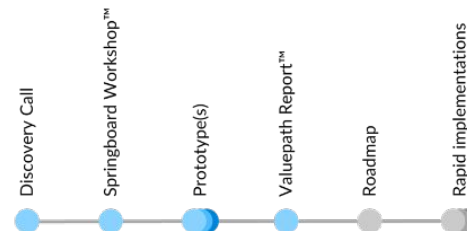
moto

NetworkRail

vision express

Onepoint & boomi

Boomi partners since 2019. 40+ data engineers trained on Boomi products. Successfully completed 4 major projects with Boomi. Recognised as Boomi's EMEA Teaming Partner of the Year 2023.

boomi
EMEA Teaming Partner Of The Year 2023
Onepoint

Onepoint Valuepath™ and the Rapid Value Method™

Onepoint Valuepath™, the upfront element of our approach, is designed to help you to **quickly prove your technical concept**. Not only that, we also help you prove value, typically within 2-6 weeks.

This approach was borne out of a business outcome focus, human centricity, diverse thinking and multi-disciplinary input, deep technical expertise, and an emphasis on delivering value rapidly.

Discovery Call — Springboard Workshop™ — Prototype(s) — Valuepath Report™ — Roadmap — Rapid implementations

Continuous innovation

We are always experimenting with the next generation of tools and techniques to validate their **ability to influence business outcomes**.

It is through continuous research and development that we are able to demonstrate the art of the achievable, without learning at the expense of our clients.

Onepoint Labs

25

Building a Data Hub in 8 Simple Steps

AI *in* the Boomi platform

vs.

Using the Boomi platform *for* AI

Left circle (AI in the Boomi platform):
- BoomiAI
- Boomi Suggest
- Autonomous Foundation

Right circle (Using the Boomi platform for AI):
- Classification & Tagging
- RAG
- Chatbots
- Agents
- Orchestrate Model Fine tuning
- Image Generation
- Synthetic Dataset Generation
- BoomiAI
- Boomi Suggest
- Autonomous Foundation
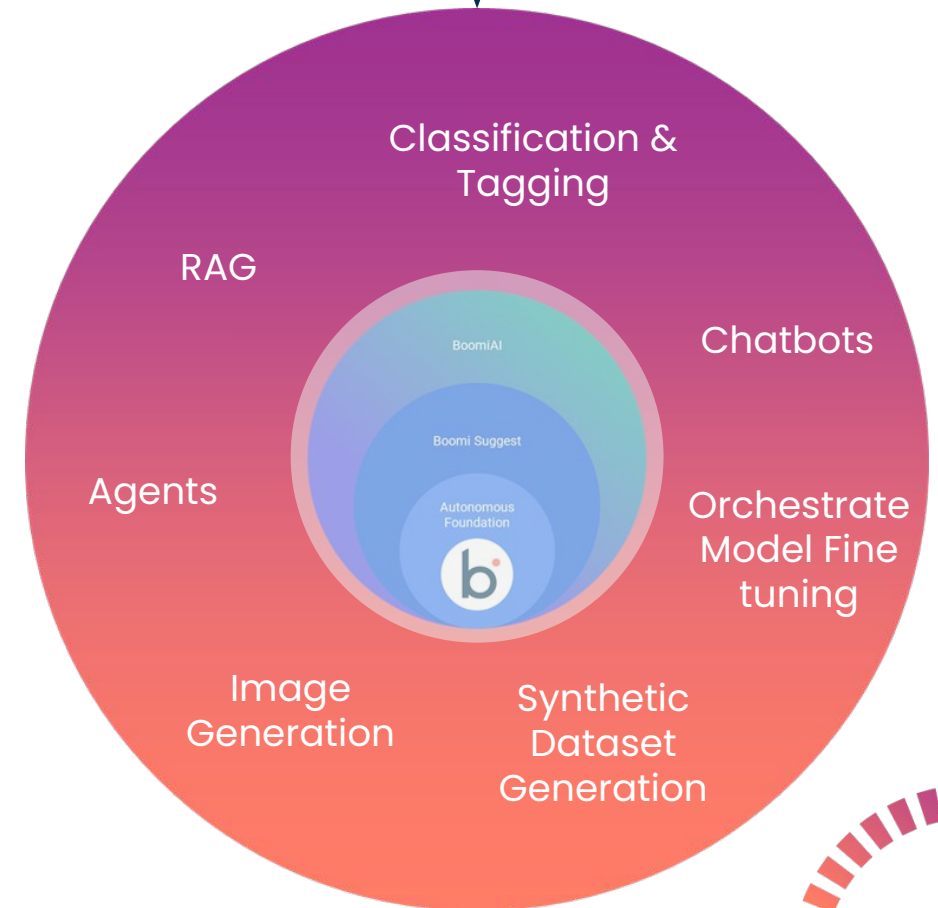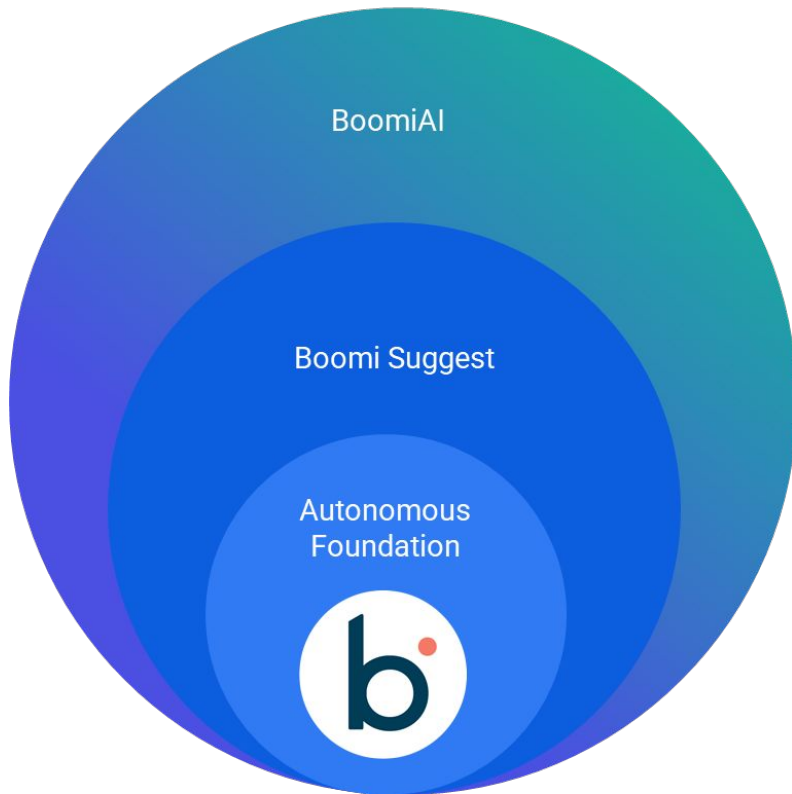
Six Tenets of AI Readiness

1 Set Clear Goals
From implementation to stewardship, knowing what you want and what you're capable of doing with AI is essential

2 Know Your Processes
Document, catalog, create, manage, and maintain workflows and business processes

3 Know Your Data
Understand the value of your data, where the data lives, and how it's generated, refined, secured, and governed

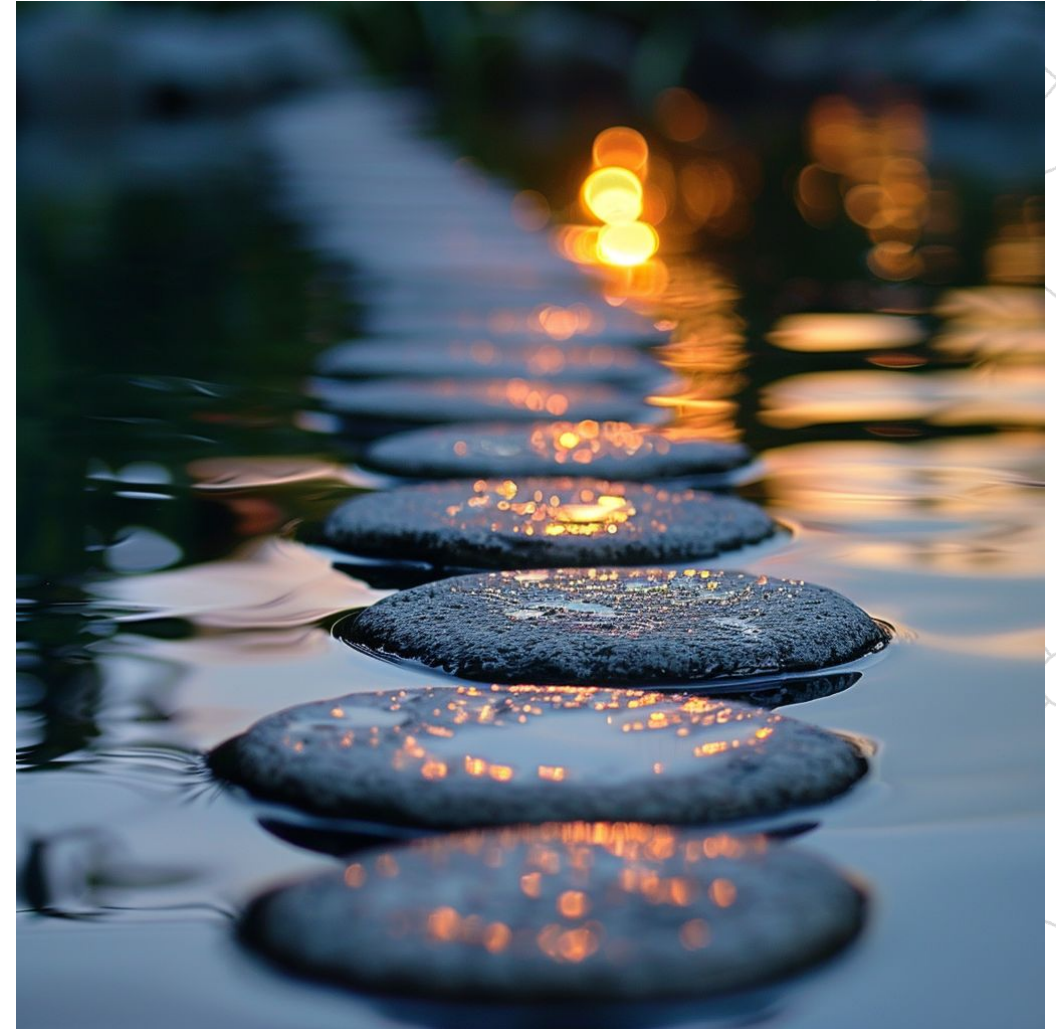**4 Align and Be Accountable**
Stakeholders should understand expectations and work together as good stewards of data and processes

5 Prioritize Thoughtfully
Establish and balance priorities critical to the business, corporate values, and societal impact

6 Automate with Intention
Implementing AI without understanding your processes or data could lead to inefficiency and will introduce risk

AI hype vs. AI adoption in enterprise

Mass Adoption

Focus of Hype

PHASE 1

PHASE 2

PHASE 3

PHASE 4

Before 2017

2017 - 2022

2020-2023

2023 onwards

- Analytical AI based on **small models**
- **Limited functionality & data**

- **Superior data quality, faster time to train & larger data ingestion** capabilities
- **Inaccessible and expensive** models

- **New and free-to-use foundation models**
- Computer capabilities **overtaking human performance**

- Increasingly **cheaper, efficient and easier to deploy models**
- **Explosion of generative AI applications**

Adapted from Insight Partners report

Master Data Management – Best Practice

Closing the loop – SYNCHRONISING master data across operational systems

Solar Energy
Generator
160 Solar Farms

quickbase
Site and Asset (API)
Project (API)

Azure Active Directory
People (API)

SAP SuccessFactors
People (API)

monday.com
IT Systems (API)

OneTrust
PRIVACY, SECURITY & TRUST
IT Systems (API)

Legal Entity (API)
Shareholder (API)

Boomi Integration

Ingest

Clean
Match
Merge

Synchronise

Site
People
Project
IT System
Legal Entity

Boomi Master
Data Hub

Quarantined
Records

Data Stewards

SFTP

People

People

People
Project

site
People
Projects

melon
powering CORPORATE TRAVELLER

SAP Concur

Culture Amp

GIS
GEOGRAPHICAL INFORMATION SYSTEMS

IFS

Microsoft Azure
Blob Storage