

Enlighten[®]

Is your data quality good enough to support your business initiatives?

www.innovativesystems.com

Enlighten[®]

Introduction

Michael Ott

Senior Vice President

www.innovativesystems.com

Application systems quickly become landfills of data



Is your data quality good enough to support your business initiatives?

You may already have data quality tools.

Are they able to deliver the quality necessary to support your new business initiatives?

Existing data quality often doesn't support new initiatives, such as customer engagement, digital transformation, compliance and analytics.



Today's session will highlight ...



Advisory Services



**Auditing
Capabilities**



DQaaS

Advisory Services



- ✓ Provides a broader view of the DQ initiative
- ✓ Ties in the people and process aspects
- ✓ Examines various components (e.g., data elements, quality of data, controls) and whether they support the business requirements
- ✓ Provides the foundation for a successful business initiative

Auditing Capabilities



Innovative Systems can quickly and accurately audit data quality early on and in conjunction with Advisory Services



Source Data

- ✓ Data to be migrated into an existing system or used to create a new target system
- ✓ Comes from various sources with different field configurations and quality standards
- ✓ Target systems are fed clean data



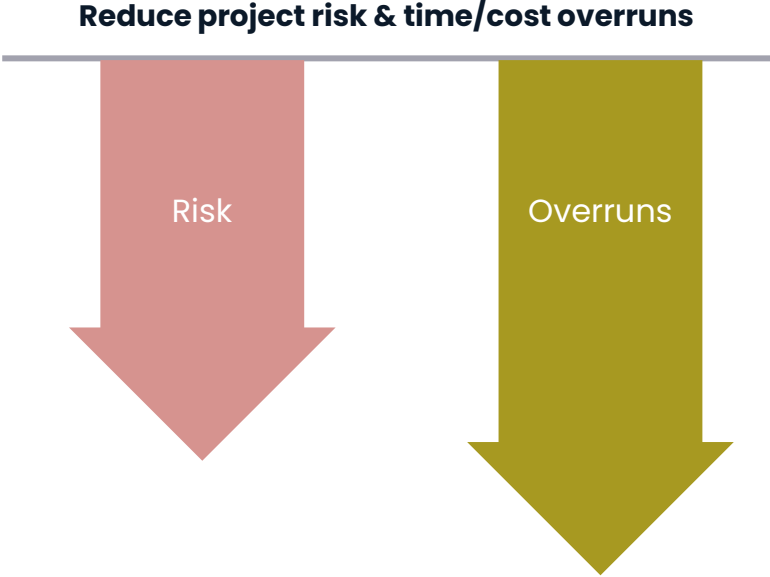
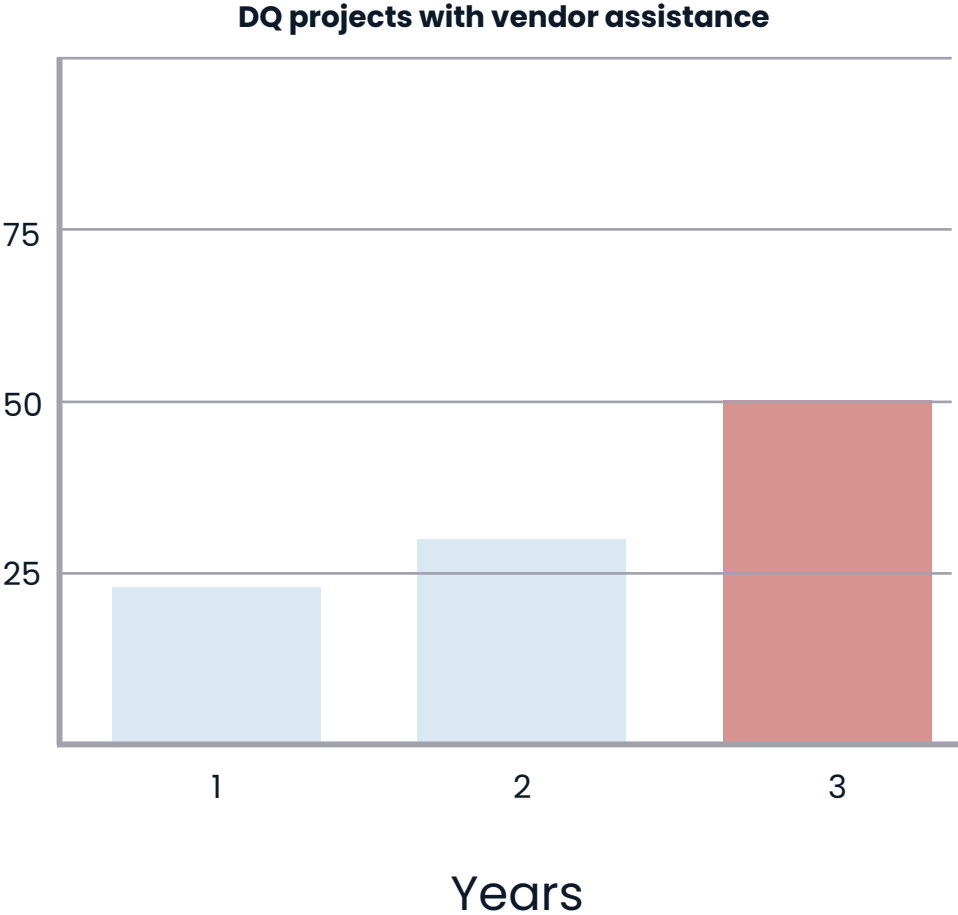
Target Data

- ✓ Data in the target system that has degraded over time from user interactions
- ✓ Monitor and maintain data quality

DQaaS



% of projects using vendor assistance



Our dictionaries have the definitions

- ✓ Creating customer-centric databases for 50+ years
- ✓ Dictionaries contain millions of words, phrases, and patterns from billions of records reviewed
- ✓ Evolving technology - 14th generation
- ✓ Users don't have to test or tune
- ✓ Processes data faster and more accurately

We don't use weighted field scoring

Threshold
80%

Match 1: 85% and a True Match

Charles	Taylor	Male	19480128	Liberia	321-45-9876	SCORE 85
Charles	Taylor		19480128	USA	321-45-9876	

Match 2: 85% and a False Match

Charles	Taylor	Male	19480128	Liberia	321-45-9876	SCORE 85
Charles	T	Male	19480128	Liberia	321-79-8990	

- ✓ Other products assign a percentage at the record level
- ✓ Anything above the threshold is considered a duplicate
- ✓ No transparency
- ✓ No fine tuning
- ✓ Over-matching or under-matching

Percentage based matching just doesn't work!

We use pattern-based matching and AI

First Name
Last Name
Title
Date of Birth

Each field pairing is compared and assigned an attribute

Equal
Close
No Match
Blank

Each pair is assigned a match string showing the likelihood of being a duplicate



Our AI engine mimics how humans process data to determine a match

Automatic Match
Review
No Link

Match strings determine if pairs are automatically matched or reviewed

Critical

Skip

Users can prioritize at the field level for total control

We use pattern-based matching and AI

Record	First Name	Additional Names	Last Name	Street	ID	Differences
Client	A	-	Taylor	Regent	142-99	
Potential link	Agnus	Reeves	Taylor	Regent	101-73	
Relationship	Close	Blank	Equal	Equal	Not Equal	Significant

Differences

Minimal



Result

Automatic Match/Link



Moderate



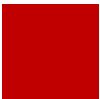
Review



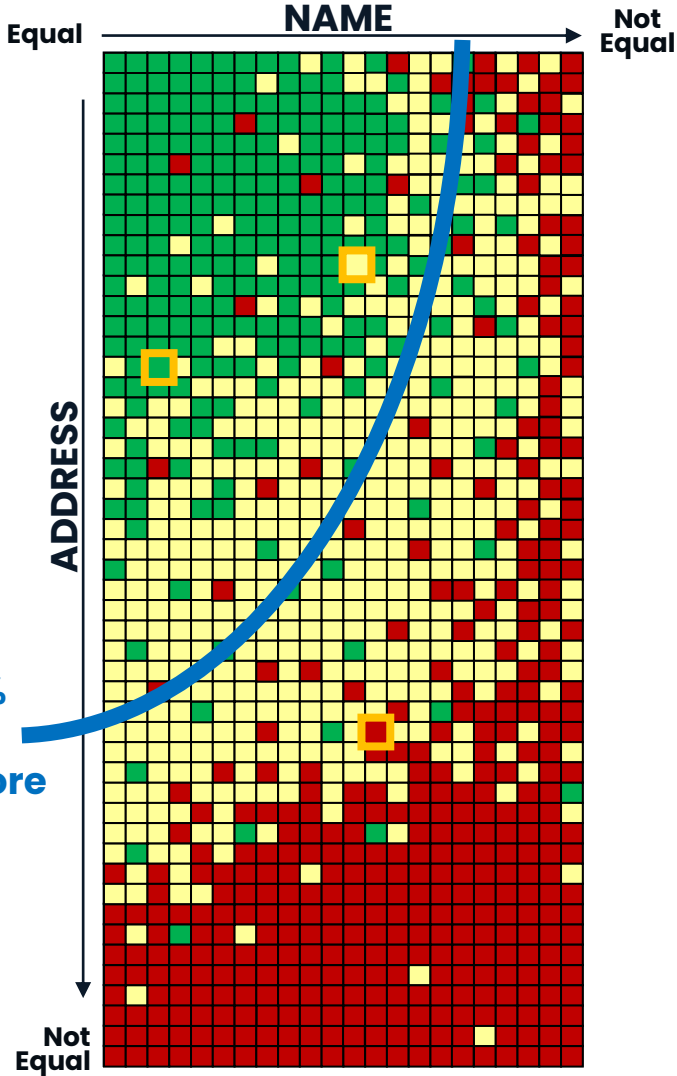
Significant



No Link



Cutoff of 80% means an acceptable score



We use natural language processing

Grammars

- ✓ Language rules
- ✓ Enable processing against any alphabet in any language
- ✓ Identify the meaning and order of data
- ✓ Can be added or changed without changing the core product



Enlighten[®]

Capabilities Demo

Hector Cordova

Director, Professional Services & Consulting

www.innovativesystems.com

Enlighten[®]

Is your data quality good enough to support your business initiatives?

Enlighten[®] Data Quality Services Engagement

www.innovativesystems.com

Data Quality Services Engagement

Case study: CRM system creation



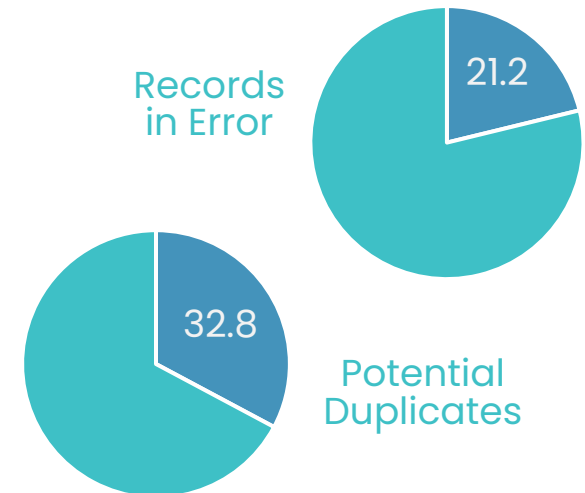
- **AUTOMAKER** requested DQ services to populate a new CRM system
- The project was conducted on **11.849 Million records** from **3 originating sources of data:**
 - CAR SALES: 6.695M records
 - CREDIT SALES: 4.193M records
 - ACQUISITION: 0.961M records
- From an **audit** perspective, the project focused on two key areas:
 - **Data Errors** (exposes potential hidden risks):
 - Multiple names on Name line
 - Hidden names on Address lines
 - Incorrect Customer Type on input
 - Invalid (default or dummy) or missing key data elements
 - Standardization/consistency errors (name, DOB, address, postal code)
 - **Duplication** (causes wasted effort, fragmented profiles)
- The data was **processed** by Enlighten and returned to the client, ready to be uploaded to the CRM system. **Prevention** measures were also put in place for incremental adds and updates.

Data Quality Services Engagement

Case study: CRM system creation



- **AUTOMAKER** requested DQ services to populate a new CRM system
- The project was conducted on **11.849 Million records** from **3 originating sources of data:**
 - CAR SALES: 6.695M records
 - CREDIT SALES: 4.193M records
 - ACQUISITION: 0.961M records
- From an **audit** perspective, the project focused on two key areas:
 - **Data Errors** (exposes potential hidden risks):
 - Multiple names on Name line
 - Hidden names on Address lines
 - Incorrect Customer Type on input
 - Invalid (default or dummy) or missing key data elements
 - Standardization/consistency errors (name, DOB, address, postal code)
 - **Duplication** (causes wasted effort, fragmented profiles)
- The data was **processed** by Enlighten and returned to the client, ready to be uploaded to the CRM system. **Prevention** measures were also put in place for incremental adds and updates.



Data Quality Service Engagement

Supporting metrics – Data Quality dimensions



Completeness

Understand what fields are “really” available



Uniqueness

Understand key attributes, unique set of values



Consistency

Ensure fields are represented in a consistent format



Accuracy

Customer type, gender against name, address



Validity

Default values, length, range, patterns



Timeliness

Active vs inactive, outdated sources of data

Enlighten[®]

Enlighten[®] Data Quality

What sets Enlighten apart?

Dictionary

- ✓ Millions of words, phrases, and patterns
- ✓ 14th generation
- ✓ Finding and correcting data is faster and more accurate
- ✓ No continual testing and tuning

Pattern Matching

- ✓ Greater accuracy than percentage-based
- ✓ Maintains field-level matches for precision
- ✓ One match string for each use case
- ✓ Total control over priorities and ranking

Language Grammars

- ✓ Linguistic-based grammars for each language
- ✓ Universal grammar for any language
- ✓ Processes data based on the language used
- ✓ No need for custom coding

What sets Innovative Systems apart?

- ✓ **Experience:** deep knowledge of working with data
- ✓ **Technology:** recruit from top universities; leader in computational linguistics
- ✓ **Analysts:** recognized leader in technology and customer satisfaction
- ✓ **Flexibility:**
 - ✓ Software: On-premise, SaaS/Cloud, Hybrid, DQaaS, Ad Hoc service projects
 - ✓ Train you or do the work for you
- ✓ **Volume:** Global data centers; process 270 billion+ records annually; meet privacy & GDPR requirements
- ✓ **Access:** 24/7 support
- ✓ **Talent:** Timely access to staff with deep technical expertise

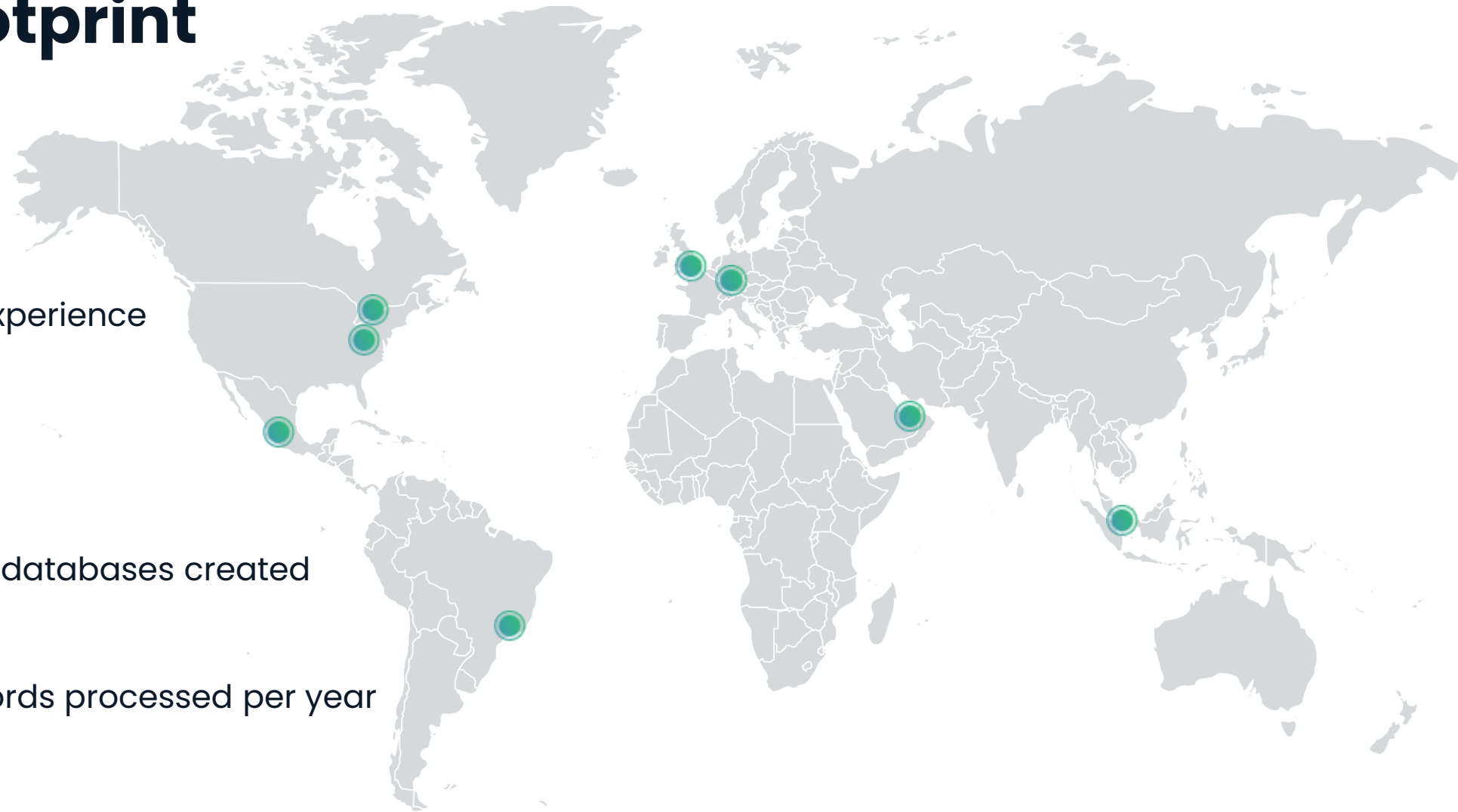
Global Footprint

50+ Years of experience

65+ Countries

1,000s Customer databases created

270+ Billion records processed per year



Locations in: Pittsburgh | London | Dubai | Frankfurt | Mexico City | São Paulo | Singapore | Toronto

Enlighten[®]

Questions?

mott@innovativesystems.com

www.innovativesystems.com