

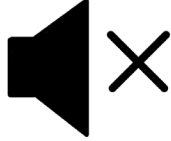
# Configuring Subject Registry and Generating Subject Access Reports using Data Privacy Management

Ramesh Swaminathan

Shishir Rai

Sheetal Modi

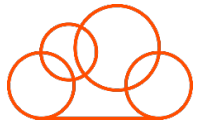
# Housekeeping Tips



- Today's Webinar is scheduled for **1 hour**
- The session will include a webcast and then your questions will be answered live at the end of the presentation
- All dial-in participants will be muted to enable the speakers to present without interruption
- Questions can be submitted to "All Panelists" via the **Q&A option** and we will respond at the end of the presentation
- The webinar is **being recorded** and will be available to view on our **INFASupport YouTube channel** and **Success Portal**. The link will be emailed as well.
- Please take time to complete the **post-webinar survey** and provide your feedback and suggestions for upcoming topics.

# Feature Rich Success Portal

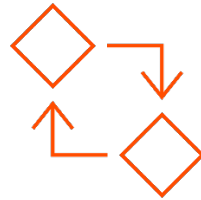
---



Bootstrap trial and  
POC Customers



Enriched Customer  
Onboarding  
experience



Product Learning  
Paths and Weekly  
Expert Sessions



Informatica  
Concierge with  
Chatbot integrations



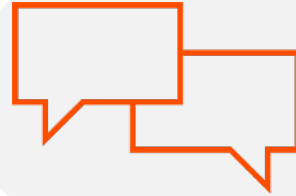
Tailored training and  
content  
recommendations

# More Information



## Success Portal

<https://success.informatica.com>



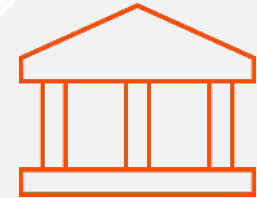
## Communities & Support

<https://network.informatica.com>



## Documentation

<https://docs.informatica.com>



## University

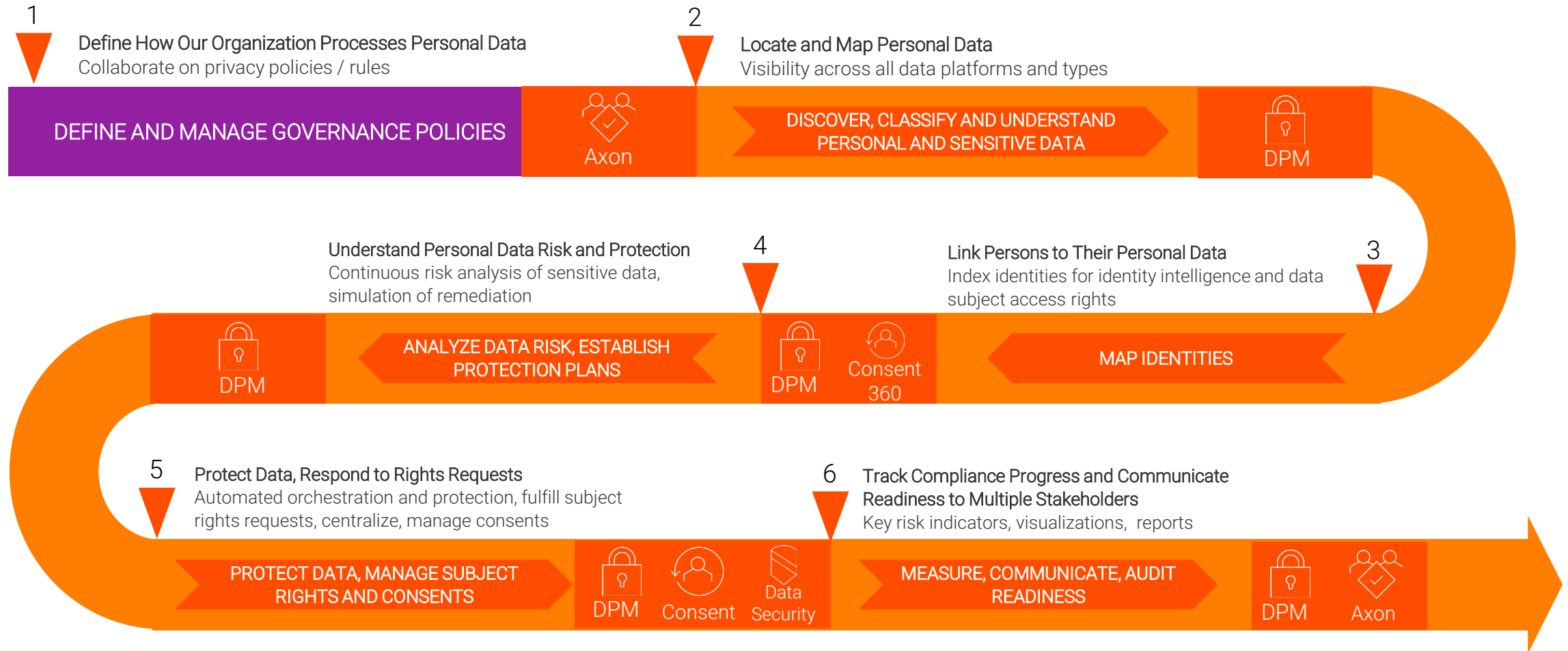
<https://www.informatica.com/in/services-and-training/informatica-university.html>

# Safe Harbor

The information being provided today is for informational purposes only. The development, release, and timing of any Informatica product or functionality described today remain at the sole discretion of Informatica and should not be relied upon in making a purchasing decision.

Statements made today are based on currently available information, which is subject to change. Such statements should not be relied upon as a representation, warranty or commitment to deliver specific products or functionality in the future.

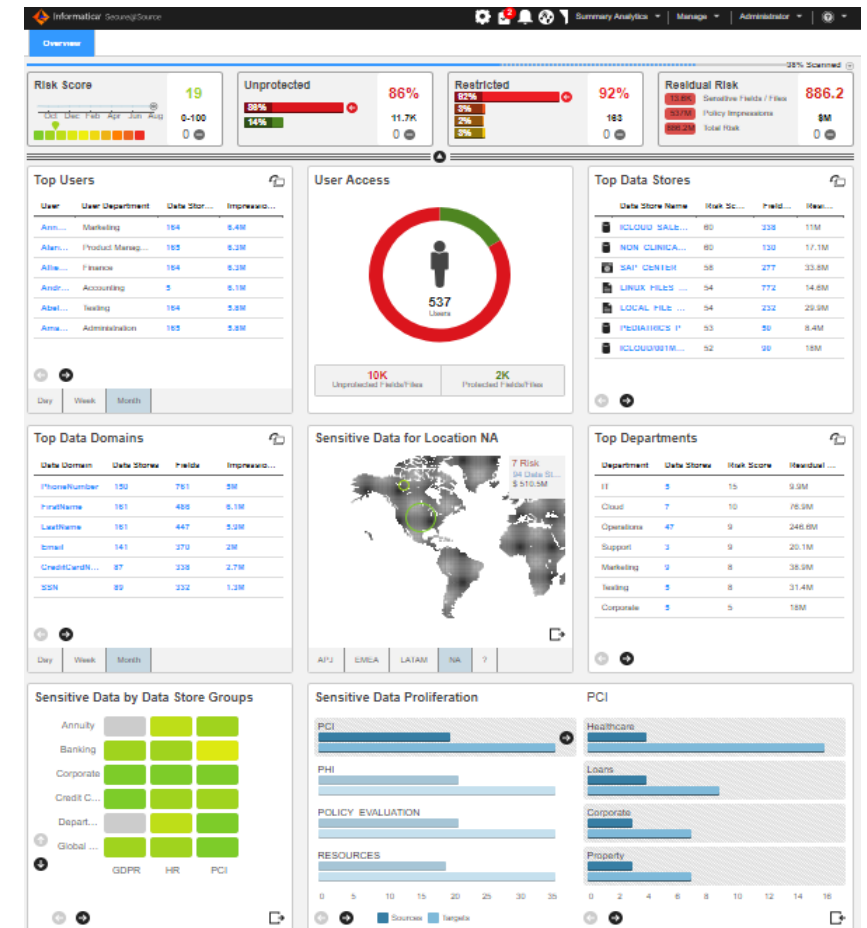
# Data Governance Framework for Privacy Regulations



# Data Privacy Management

## ➤ Enterprise-wide data privacy and security intelligence

- Sensitive data discovery and classification
- Data subject registry
- Sensitive data proliferation analysis
- User activity on sensitive data
- Multi-factor risk scoring
- Risk rankings at enterprise, LOB, and department level
- Anomaly detection based on user activity
- Sensitive data policy-based alerting
- Data protection simulation and planning
- Workflow for manual or automated data protection and risk remediation



# Subject Registry Flow



## Entity Configuration for Each Subject Type

- Entity Field types
- Match Configurations for Index creation
- Golden data sources
- Search Configuration
- Match Rules

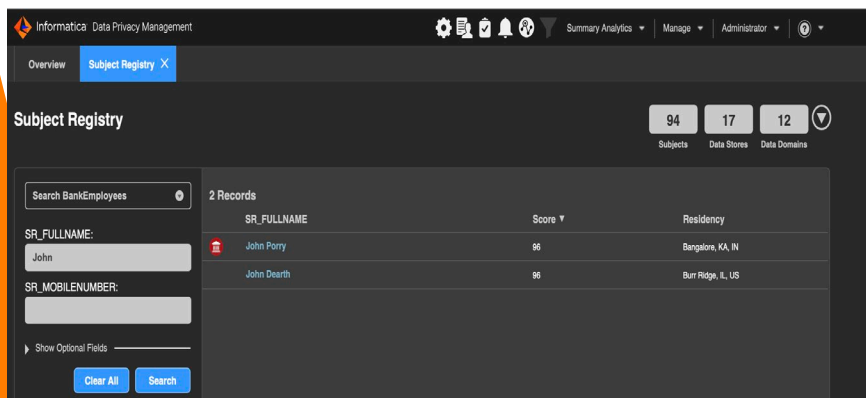


## Data Extraction Sqls for DSAR reports



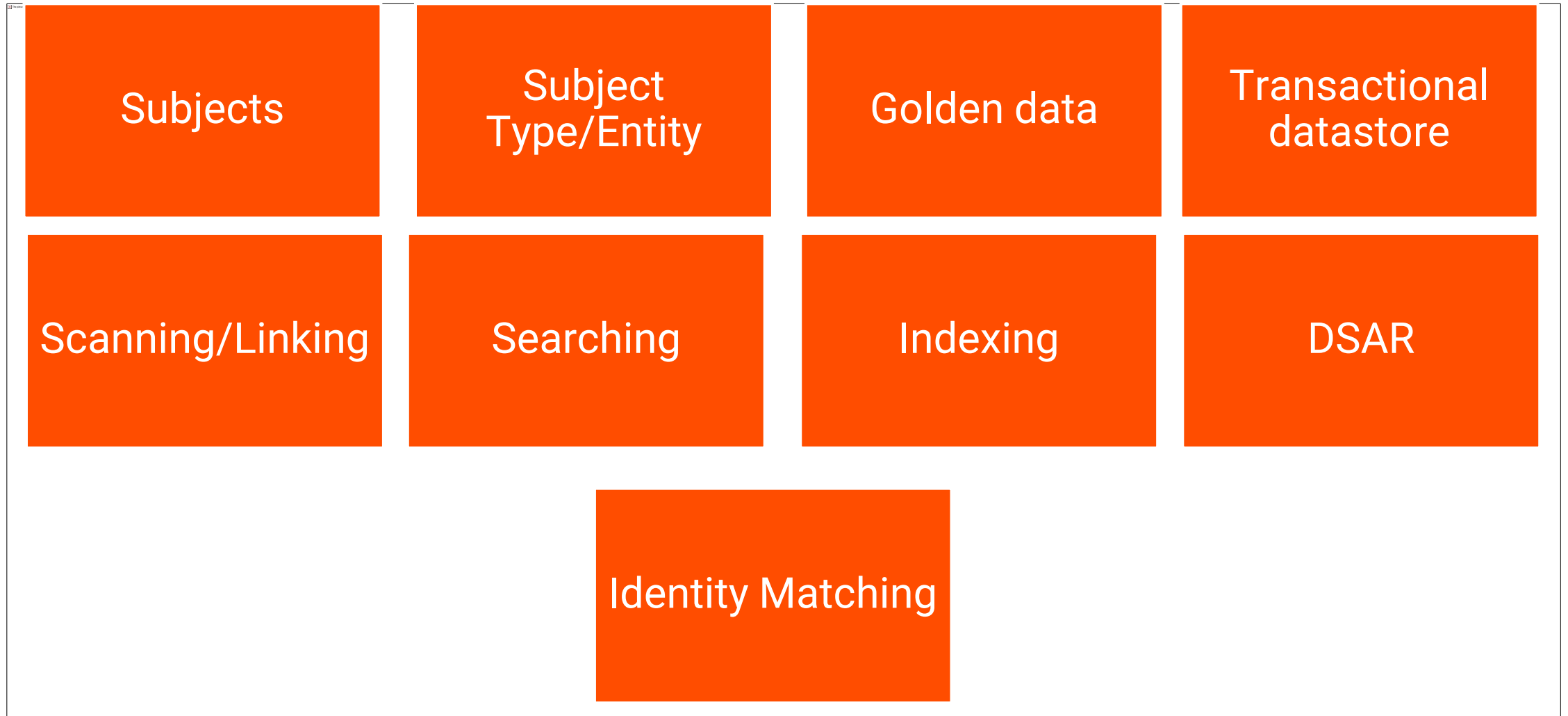
## Subject Registry Scans

- Profile Scans
- Subject Index Scans
- Linkage Scan





# Important terminologies

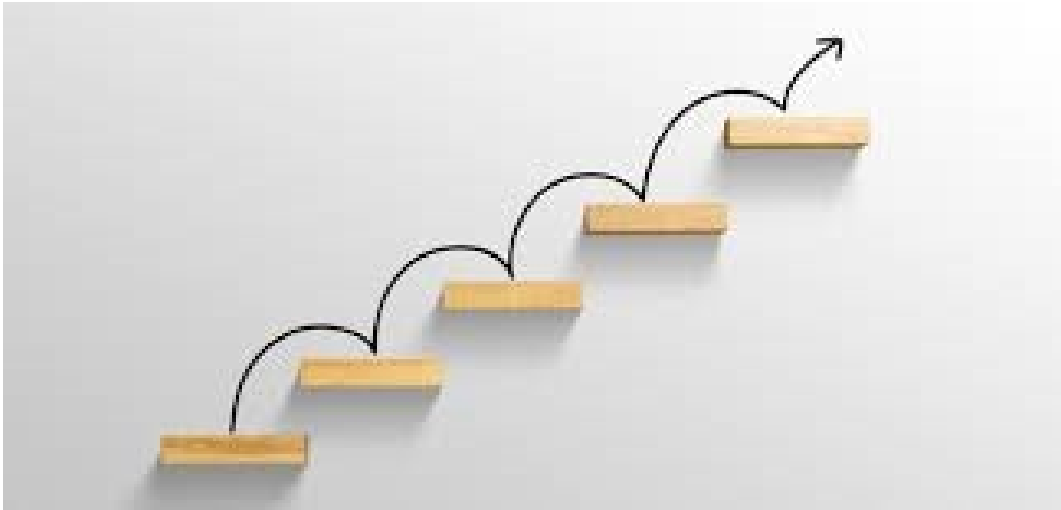


# Golden Data Sources – Single vs Multiple

Golden Data Source – A term borrowed from MDM world. A data source which is a source of truth that contains Master data about a specific subject type eg, Customer data, Employee data etc

- DPM supports multiple Golden data sources for Subject registry configuration
  - Configure a single data source as golden if you have master data of the subject you are configuring is in a single source
  - Configure multiple data sources as golden if master data about of the subject is distributed across multiple sources
  - You do NOT need a MDM solution implemented to use Subject Registry in DPM

# Incremental Golden Scans



- Incremental golden scans need to be run so that new subject records are inserted into the registry
- New records need to be created as new subjects are added into your system or existing data need to be modified
- Incremental scans enable you to run scans only on newly added or modified records

# Entity Fields



Fields in SQL must be equal to Fields in entity



Any field which will be used in scan/search/DSAR should be part of entity fields.



Fields which has length > 0 will be encrypted and stored in our persistent layer.



Fields are truncated by the field length.



Store only those fields ( minimum set) which can uniquely identify a subject.

# Match Configuration

01

Index  
Configuration

02

Search  
Configuration

03

Match Rules

# Index Configuration

---

Used to identify the fields for generating indexes.

---

Indexes are generated only when running a golden scan. So if we need to change index, we need to run golden scan again

---

Indexes are useful to improve both scan and search performance .

---

Index Type – Fuzzy and Exact.

---

# Fuzzy Index vs Exact Index

- Use Fuzzy index where matching a certain pattern/characteristics are desirable rather than exact values. These characteristics might have arisen due to data corruption like wrong spelling , using nicknames , using shorthand notation etc.
- Most common use case is matching Names , Addresses etc.
- Use Exact index where no value is acceptable other than exact matches like employee ID, customer ID, phone number, SSN etc.
- One can use exact index for matching names during scan job and fuzzy indexes while searching on UI

# Ideal candidates for exact indexing(Scan)



Uniquely identify the records like Ids.



Present in all rows/data.



If above two criteria satisfied, then no other fields needed for linking records and will be highly performant.



If a field is non-unique then it can be used for exact indexing provided user has mentioned extra fields for matching in match rules to break the tie ( same index values but different records).



# Ideal candidates for fuzzy indexing(scan)



Should be present in all records with no null values.



String values are preferred over numerical.



If numerical, better use exact index. As fuzziness of numeric value doesn't make much sense.

# Search Configuration

---

Search configuration is used both in scan and search.

---

In scan, it is used in Transactional scans and Multiple golden scans.

---

The fields used in search configuration should be a subset of index configuration.

```
"MatchConfigurations": [{  
  "MatchConfigurationName":  
  "MatchConfiguration1",  
  "IndexConfiguration": {  
    "FieldName": "FULLNAME_E",  
    "IndexType": "Fuzzy",  
    "IndexPurpose": "Person_Name",  
    "IndexLevel": "Standard"  
  },  
  "SearchConfiguration": {  
    "FieldName": "FULLNAME_E",  
    "SearchType": "Fuzzy",  
    "SearchPurpose": "Person_Name",  
    "SearchLevel": "Typical"  
  },  
}
```

# Match Rule

---

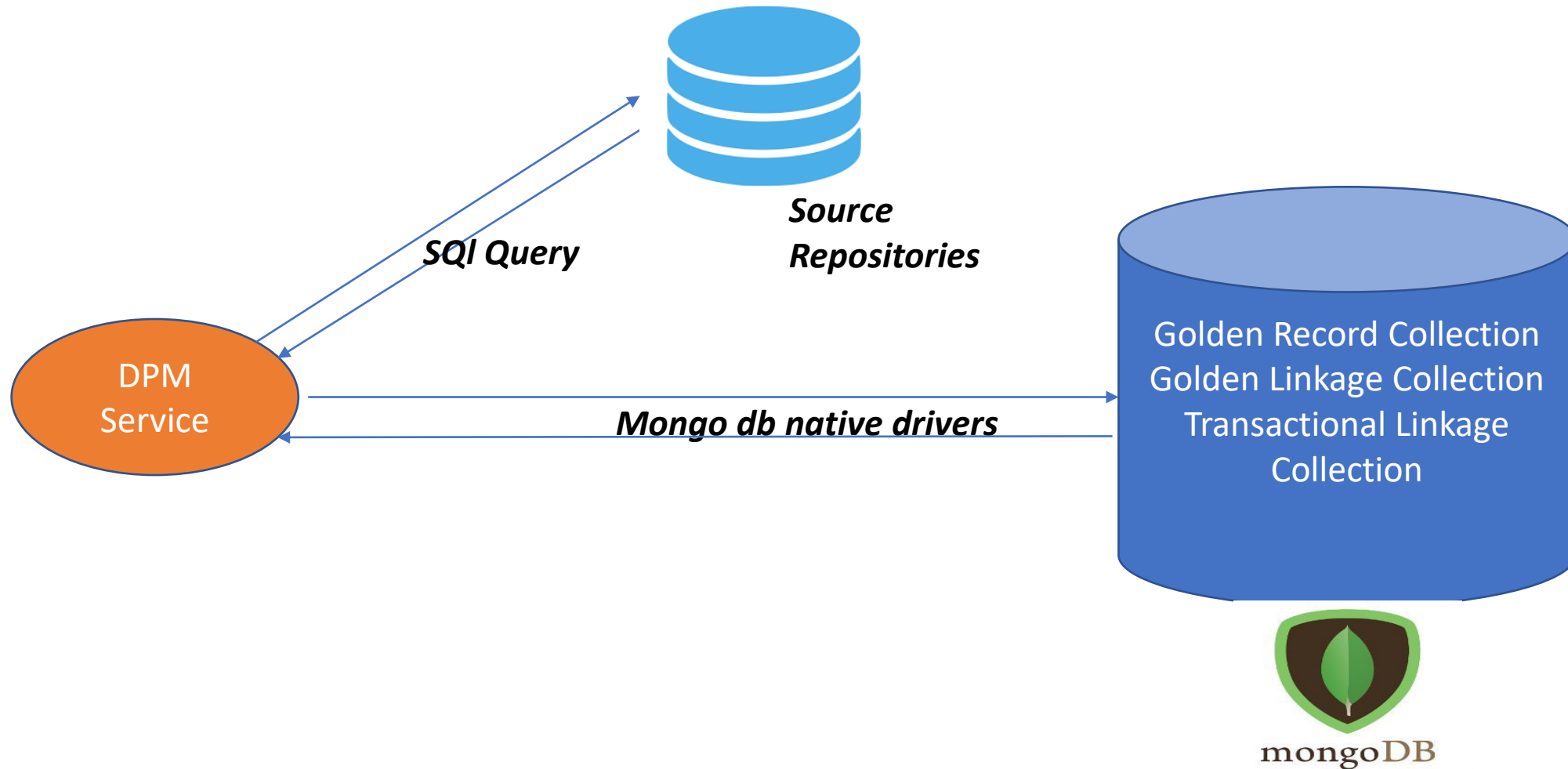
Match rule contains the logic of linking and searching the different subjects.

---

Each match rule should contain smallest subset of field which can uniquely identify a subject.

---

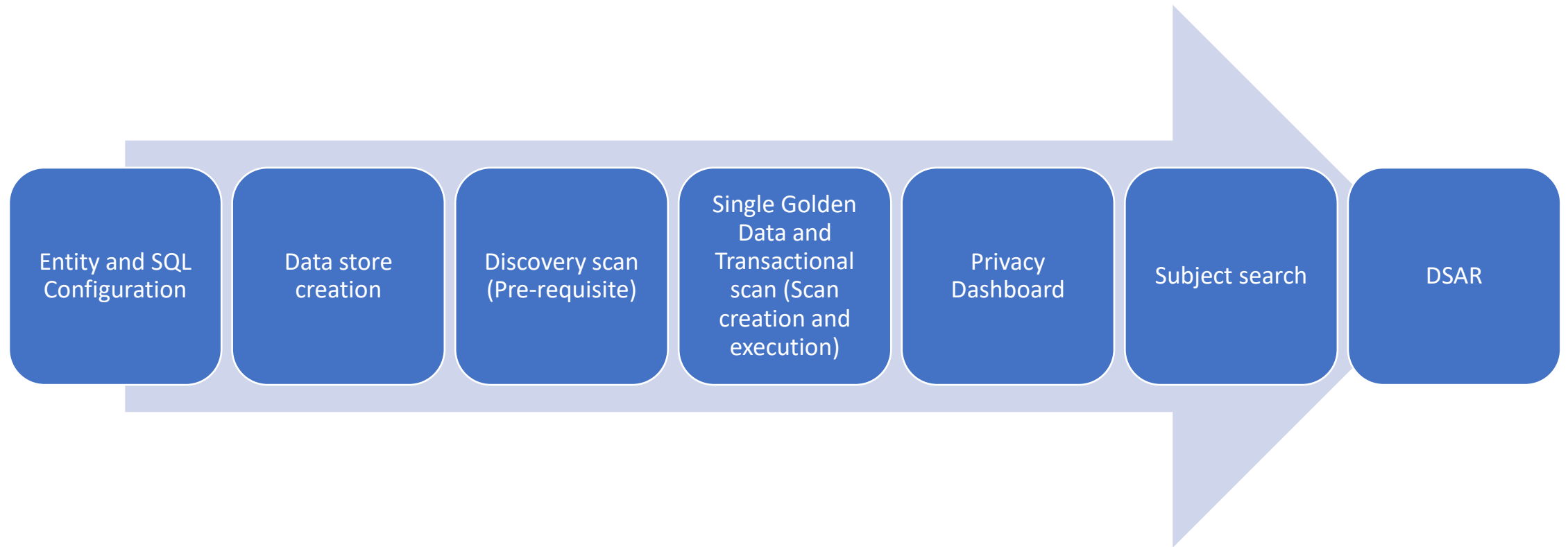
# Data flow during scans



# DEMO

- Entity Configuration
- Single Golden Data and Transactional scan
- DSAR

# DEMO FLOW



Questions?



Thank You