# Hands On With COmanage Match

COmanage™

2022 Technology Exchange

Presenters: Benn Oshrin and Laura Paglione

# Installing Match

Quick Install

# How many have Match installed?

Resources:

https://bit.ly/hands-on-with-match

# Hands On With COmanage Match

2022 Technology Exchange

Presenters: Benn Oshrin and Laura Paglione

# Introductions

Why are you spending your Monday morning here?

- Name
- Affiliation
- Why COmanage Match?

# About

Some background about Match

- The COmanage Project
- What is COmanage Match?

# The COmanage Project

- Open Source Project with generous ongoing support from InCommon

- GOAL: Streamlining digital lifecycle management for your populations

- Two tools
  - **COmanage Registry**: Identity enrollment & Lifecycle Management. Used as central person registry, guest management system, or collaboration hub for scholarly collaborations.

  - **COmanage Match**: Inbound Record Cross-system Matching. Can be used with Registry or standalone to complement your other systems.

# COmanage Match

v1.0.0, July 2022

- Matching across multiple authoritative systems of record (SORs)

- Heuristic-based system

- Flexible rule configuration

- Matching attributes defined by you

- Use w/Registry or standalone

- A full API

# Example Use Cases

- Information about the same person
- Tricky matching
  - Fuzzy matching
  - Same, but different (alternate forms of the same information)
  - Information stored in different fields
- Matching legacy information to current systems
- Non-person data matching

# Terminology

- **System of Record (SoR)** - authoritative source of information for data set

- **Record** - a structured set Attributes values that represent a Data Subject

- **Data Subject** - The person, place, or thing described by a Record in a SoR

- **Attributes** - Data fields used to describe the Data Subject in a SoR

- **Matching Rules** - A set of instructions used to compare Attribute values in Records to determine if they describe the same Data Subject.

# About Identity Matching

- Coordinated vs independent attributes

- Synchronous vs asynchronous match resolution

- Implementation at registry vs standalone

- Match before registry vs match at registry

# Coordinated vs independent attributes

What attribute values are compared to during matching.

**Coordinated Attributes**

- "Golden" attributes

- All SoRs agree to use these attributes

- Matching is done against the golden set

**Independent Attributes**

- Each SoR is authoritative for its attributes

- Matching may be against a canonical representation of attributes

- Alternatively, matching may be against each SoR's representation of attributes

# Synchronous vs asynchronous resolution

When matches are resolved if there is a partial/potential match.

**Synchronous Match Resolution**

- Resolved by the person providing the data

- Resolution required before proceeding

**Asynchronous Match Resolution**

- Resolved by an "administrator"

- Notification sent for resolution later

- Must be used for batch resolution

# Implementation at registry vs standalone

How the matching service is integrated into your workflow.

**Implementation at Registry**

- Identity matching is integrated into the person registry

- COmanage Registry and other tools have integrated Match

**Standalone Implementation**

- The Match engine independently provides information about matches

- The results of matches can be viewed or consumed by your systems

# Match before registry vs match at registry

When is matching done? Before or after the record is in the registry

**Match Before Registry**

- Matching before the record is added

- May happen during an enrolment flow

- Useful if you don't want the SoR record added if there is a match

**Match At Registry**

- Matching at the time of adding the record

- Information from the source is added first and then matched to aggregate it
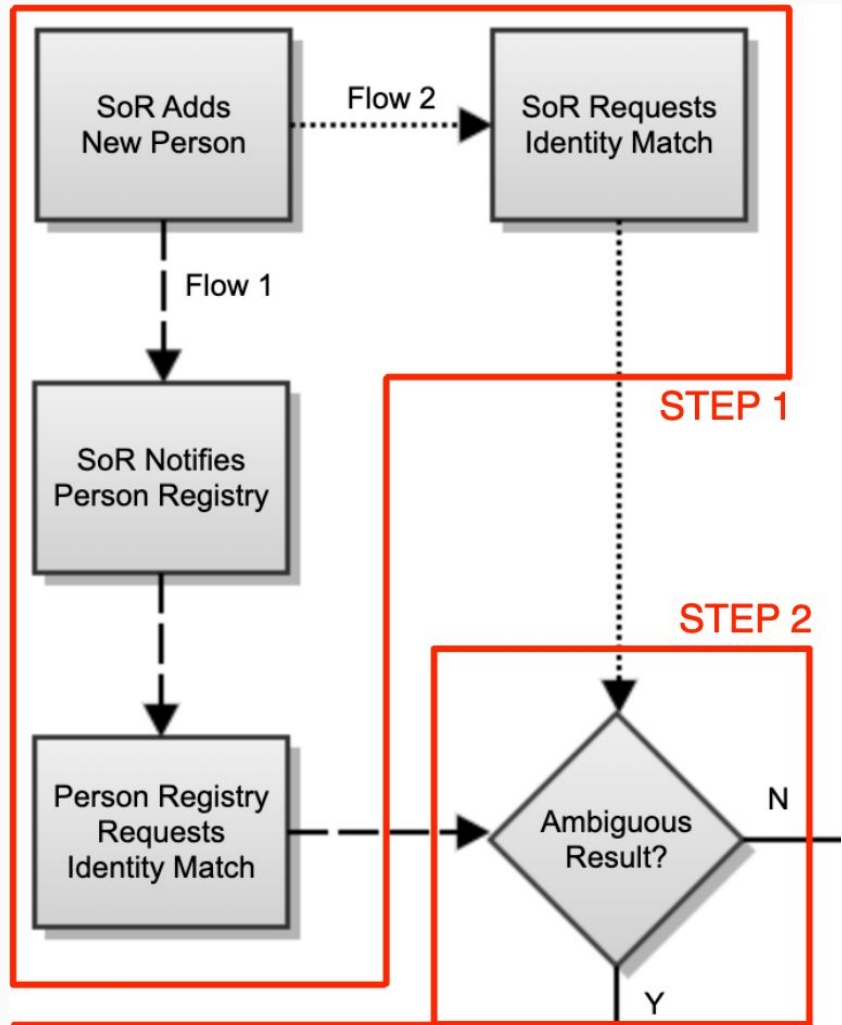
# Identity Match Process Flow
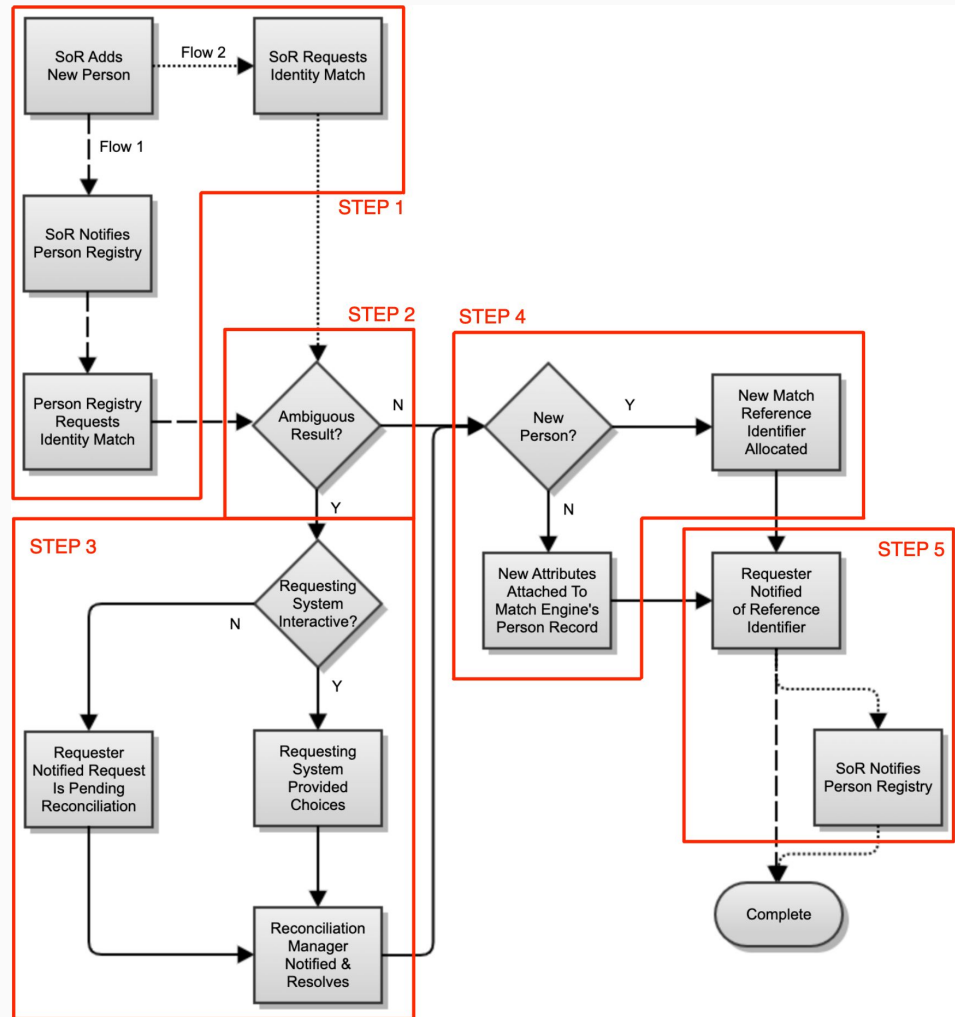
# Identity Match Process Flow

STEP 1
## New Record from a SoR is ready for Matching

FLOW 1 The System of Record provides a record for new data subject (person) to the (person) registry. The registry sends the information to Match for evaluation. This information is called an "Inbound Record".

FLOW 2 The System of Record (or other system) directly provides a record for new data subject (person) to Match for evaluation. This information is called an "Inbound Record".
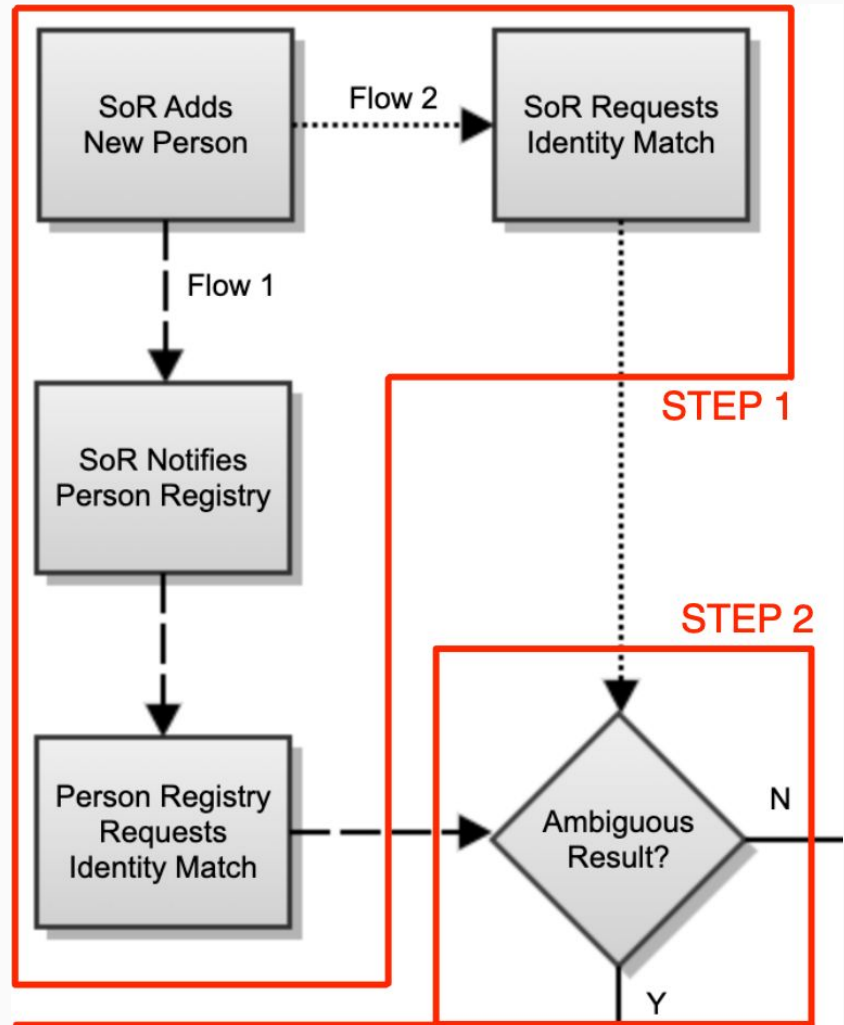
# Identity Match Process Flow

STEP 2
## Match Rules are evaluated

The Match engine searches for matching Records from the set it knows about. For each Inbound Record, the rules will produce either an ambiguous or unambiguous result. Records with ambiguous results will proceed to STEP 3; those with unambiguous results will proceed directly to STEP 4.
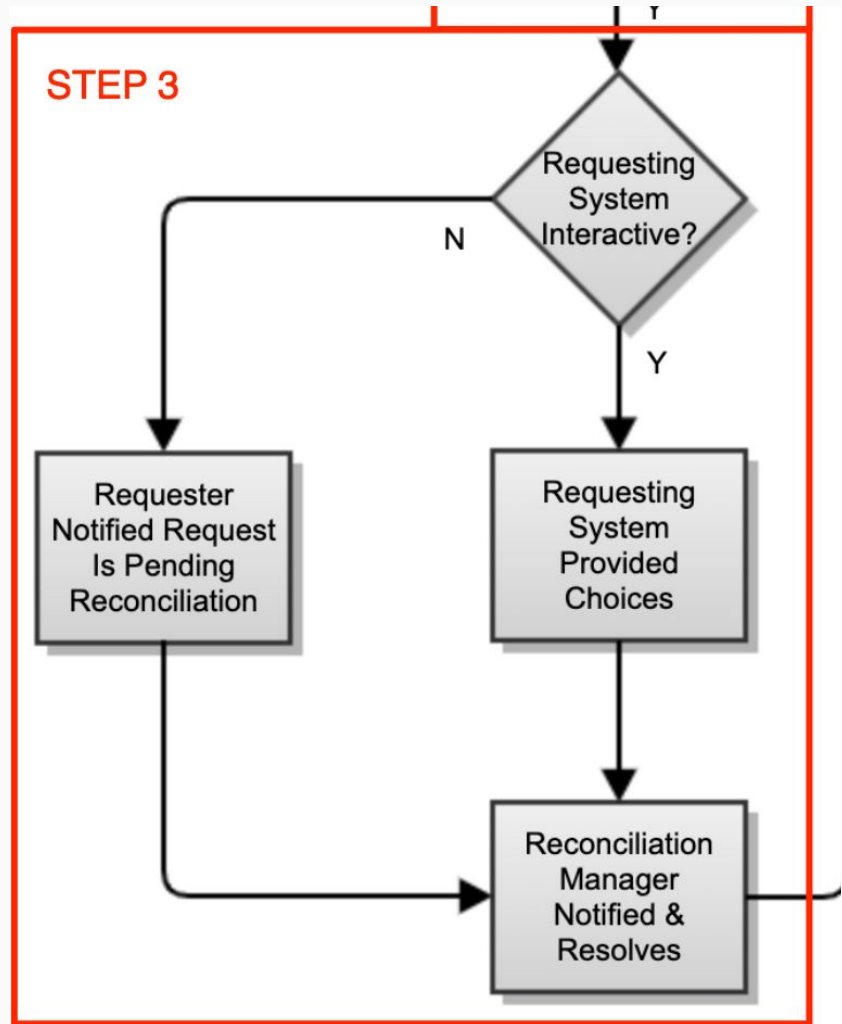
# Identity Match Process Flow

STEP 3
## Potential Matches are Resolved

If the rules make an ambiguous match, an individual needs to resolve it. Match can resolve synchronously (interactive) or asynchronously.

ASYNCHRONOUS - A notification is sent to the configured email address that a potential match needs to be resolved. An individual will resolve the match before this Record will be assigned a Match Reference ID, but this resolution doesn't have to be done right away.

SYNCHRONOUS - The system requesting the match is sent data to be presented to the individual to resolve the match right away. Note that this flow, if chosen, requires integration with a system outside of COmanage Match.
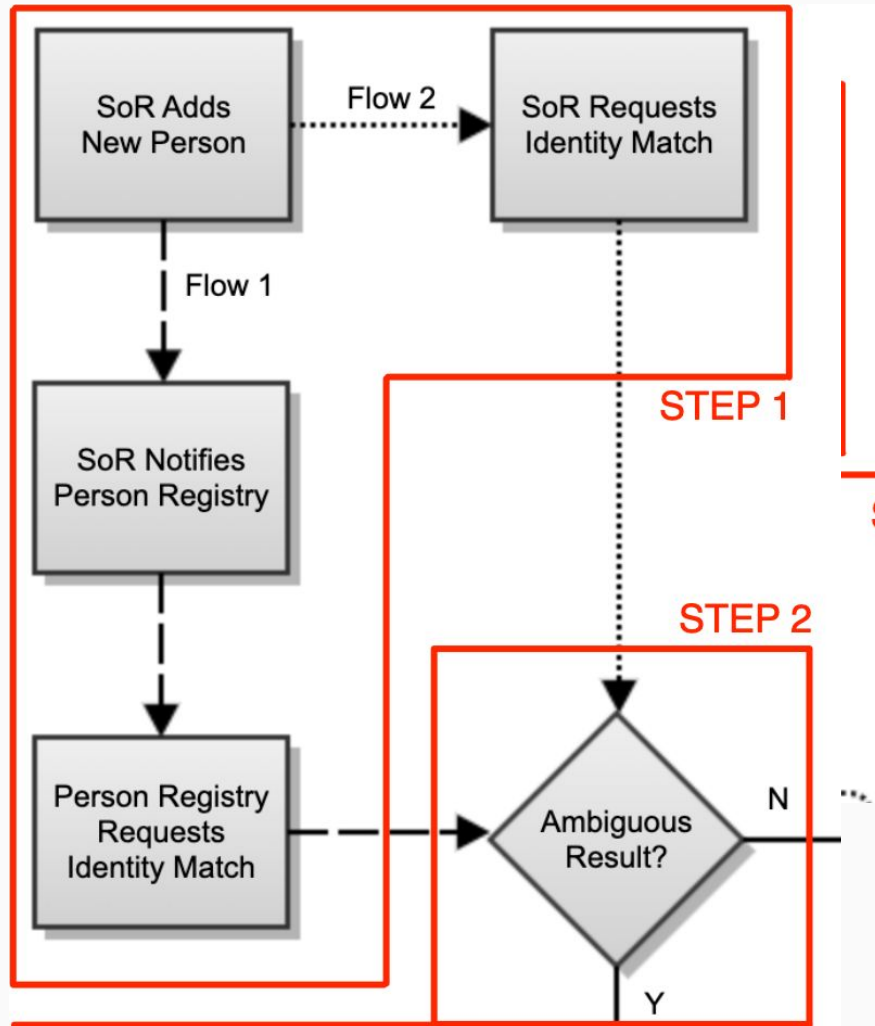
# Identity Match Process Flow

STEP 4
## Match Reference Identifier is Assigned

Based on an unambiguous result from the match rules, or a resolve match by an individual, Match will assign a match reference identifier:

NEW RECORD - if the Inbound Record represents a unique data subject (person), Match will assign it a new reference identifier.

MATCHED RECORD - if the Inbound Record is determined to represent the same data subject (person) as one already identified, Match will assign the Inbound record the reference identifier of the matched existing Record. NOTE: the records remain distinct. The presence of the same reference identifier indicates that they represent the same data subject (person).
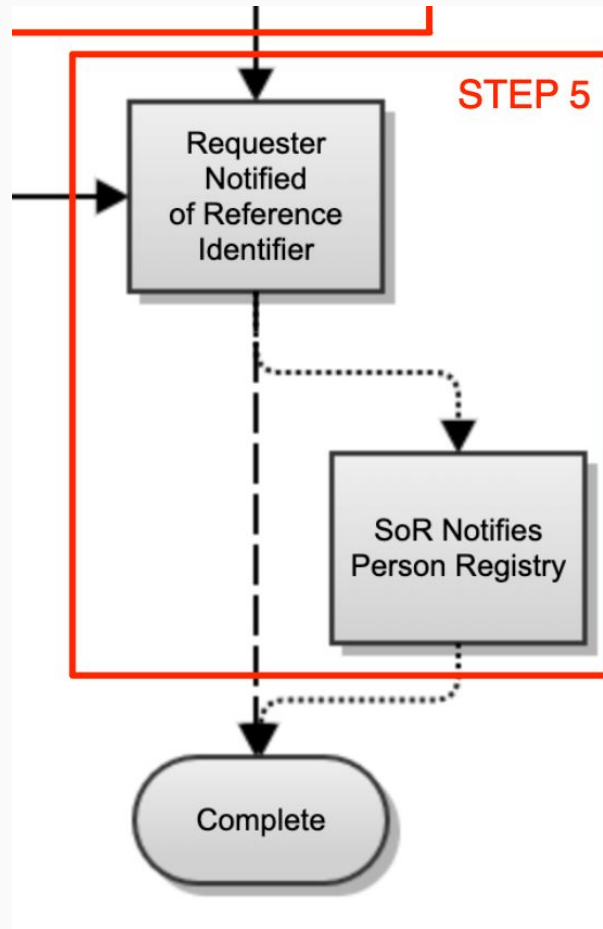
# Identity Match Process Flow

## STEP 5
### Notifications are Sent

Once a match reference identifier has been assigned to an Inbound Record, the system requesting the match analysis is sent a notification containing the reference identifier for the Record. For FLOW 1, the requesting system is the (person) registry. Match stores records that have been matched so they are available for matches to subsequent new Inbound Records.

# Identity Match Process Flow

1. New record from a SoR is ready for matching

2. Match rules are evaluated

3. Potential matches are resolved

4. Match reference identifier is assigned

5. Notifications are sent
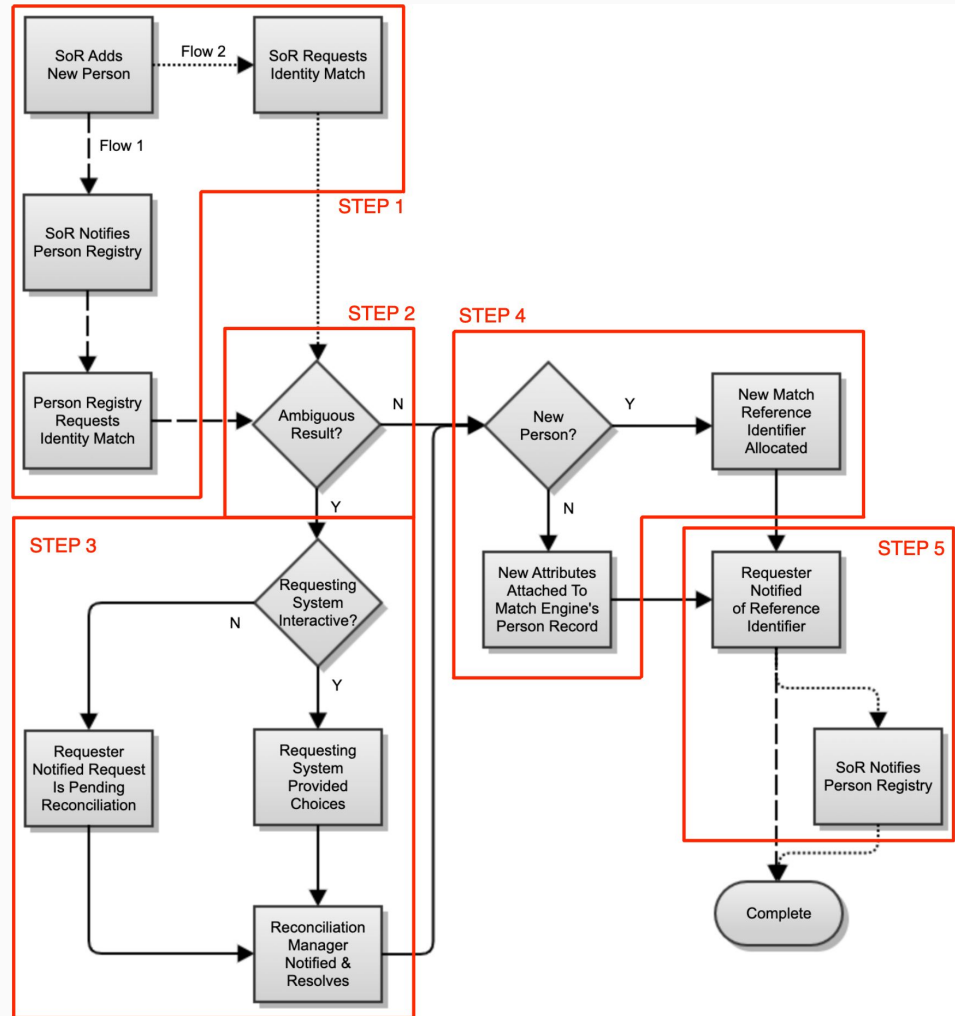
# Installing Match

Quick Install

# How many have Match installed?

Resources:

https://bit.ly/hands-on-with-match

# BREAK

# The Matchgrid

Where everything starts

# Each Match tenet is a "Matchgrid"

# Add a Matchgrid

## Add New Matchgrid

**Table Name** *

*Unique name for matchgrid, must be a valid SQL identifier (will be prefixed mg_ for actual table name)*

**Description**

**Status** *

Active

*\* Denotes Required Field*

SAVE

|  | **Demo Matchgrid for today** |
| --- | --- |
| **Table Name** | `demo` |
| **Description** | `Demo Matchgrid for use with an external registry` |
| **Status** | `Active` |

# Match Users / Permissions

# Match Users / Permissions

COmanage Match > Permissions > Edit demouser

## Edit demouser

| Username * | demouser | ⊡ |
|---|---|---|
| Permission * | Matchgrid Administrator | ⌄ |
| Matchgrid * | demo | ⌄ |

*Denotes Required Field*

**SAVE**

# Enter the Matchgrid

## COmanage Match
COmanage™

## Matchgrid Selection

Select the Matchgrid you would like to work with.

| Matchgrid | Description |
| --- | --- |
| demo | Demo Matchgrid for use with Demo Registry |

# How COmanage Match indicates a match

Match Reference ID

- Two possible formats
  - Sequence
  - UUID (type 4)

- Records that match will be assigned the same Reference ID

# Matchgrid Settings

# Matchgrid Settings

# Systems of Record

Data comes from somewhere

# Configure Systems of Record

## Matchgrid: demo

COmanage™

- 🏠 Manage
- ☰ Display
- ✓ Reconcile
- ⚙ Configure

COmanage Match > demo > Configure demo

# Configure demo

- ⚙ **Matchgrid Settings**
- ☰ **Attribute Groups**
- ⟋ **Systems of Record**
- 🔑 **API Users**
- ⇄ **Attribute Maps**
- 🔧 **Build**
- ✏ **Attributes**
- 📋 **Rules**

# Add System of Record



**Matchgrid: demo** · COmanage™

COmanage Match > demo > Configure > Systems of Record

## Systems of Record

Manage
Display
Reconcile
Configure

⊕ Add New System of Record

| Label | Resolution Mode | Action | |
|-------|-----------------|--------|--------|
| hrms | External | ✏ Edit | 🗑 Delete |
| sis | External | ✏ Edit | 🗑 Delete |

1 of 1

# Defining Systems of Record



**Matchgrid: demo** COmanage™

- Manage
- Display
- Reconcile
- Configure

COmanage Match > demo > Configure > Systems of Record > Edit hrms

## Edit hrms

| Label * | hrms |
|---|---|
| Resolution Mode * | External |
| Notification Email<br>*On potential match, notify this address of the request requiring resolution* | |

\* Denotes Required Field        **SAVE**

|  | HR System of Record | Student System of Record |
|---|---|---|
| **Label** | `hrms` | `sis` |
| **Resolution Mode** | `External` | `External` |
| **Notification Email*** | <blank> | <blank> |

*\* Notification Email is available by editing an already-created system of record*

# Attributes

Where comparisons are made

# Anatomy of an attribute

- Which attributes from the sources do you want to compare?
-

# Anatomy of an attribute

**Descriptive Information**

- Attribute Name
- Description
- API Name
- Attribute Group
- Display field in the Matchgrid?

**Matching Traits**

- Alphanumeric
- Case Sensitive
- Null Equivalents
- Search Distance
- Search Exact
- Search Substring from/for
- Attribute Map

**We'll focus on these first!**

# Attribute Configurations

Matchgrid: demo     COmanage™

COmanage Match > demo > Configure demo

## Configure demo

- 🏠 Manage
- ☰ Display
- ✅ Reconcile
- ⚙️ Configure

⚙️ Matchgrid Settings    ☰ Attribute Groups    ⚔️ Systems of Record

🔑 API Users    ⇄ Attribute Maps    🔧 Build

✏️ Attributes    📋 Rules

# Attribute Groups

COmanage Match > demo > Configure > Attribute Groups

## Attribute Groups

➕ Add New Attribute Group

| Name | Action |
|------|--------|
| official | ✏ Edit  🗑 Delete |

1 of 1

# Attribute Mappings

# Add and configure attributes



Matchgrid: demo    COmanage™

COmanage Match > demo > Configure > Attributes

## Attributes

⊕ Add New Attribute

| Name | Attribute Group | Action |
|------|-----------------|--------|
| dob | | ✏ Edit   ⧉ Duplicate   🗑 Delete |
| firstname | official | ✏ Edit   ⧉ Duplicate   🗑 Delete |
| lastname | official | ✏ Edit   ⧉ Duplicate   🗑 Delete |
| ssn | | ✏ Edit   ⧉ Duplicate   🗑 Delete |

1 of 1

Manage
Display
Reconcile
Configure

# Configure attributes

## Edit ssn

**Name** *
*Value must be a valid SQL identifier, as it will be used to construct the matchgrid column name*

ssn

**Description**

**API Name** *

identifiers:identifier/national

**Display Field in Matchgrid Index**
☐

**Alphanumeric**
☑

**Case Sensitive**
☐

**Null Equivalents**
☑

**Search Distance**
2

**Search Exact**
☑

**Search Substring From**

**Search Substring For**

**Attribute Map**

**Attribute Group**

*\* Denotes Required Field*

**SAVE**

|  | **Date of Birth** | **Email address** | **Given Name** | **Family Name** | **Identifier** |
|---|---|---|---|---|---|
| **Name** | dob | email | firstname | lastname | ssn |
| **Description** | Date of birth | Email address | Given name | Family name | National ID |
| **API Name** | dateOfBirth | emailAddresses:address/ primary | names:given | names:family | identifiers: identifier/ national |
| **Attribute Group \*** |  |  | official | official |  |
| **Display in Matchgrid?** | ☐ | ☑ | ☑ | ☑ | ☐ |

*\* Attribute Group*

"dateOfBirth":"1983-03-18",                          :[
    {
        "type":"primary",                    ational",
        "address":"pat.lee@recurrant.edu"  er":"3B902AE12DF55196"
    }
]

# Rules

What constitutes a match?

# About Match rules

- Each Matchgrid will have one or more rules
- Each rule will have a configuration for one or more attributes
- Rules are configured with a confidence mode
  - **Canonical** - definitively defines if records are matched or not
  - **Potential** - may identify matching records.

- Rules behave like searches. An attribute configuration are the settings to find the incoming attribute value in the set of records in the Matchgrid.

# Processing Match rules

## I. Process Canonical Rules

1. Each rule is evaluated
2. Processing stops when a rule finds one or more matches *(i.e., a search produces a results set)*
3. If exactly one match is found, the record is given the same Reference ID as the match
4. If more than one is found, it is treated as a potential match

## II. Process Potential Rules

1. ALL potential rules are evaluated
2. Each record that matches (items found during a search for the attributes) is a potential match
3. If processing is synchronous, the match must be resolved: select a match, or indicate no match
4. If processing is asynchronous, a notification is sent for later resolution

# Managing Rules

## Matchgrid: demo

COmanage™

- Manage
- Display
- Reconcile
- **Configure**

# Rules

**Add a rule**

**Configure a rule**

⊕ Add New Rule

| Name | Confidence Mode | Order | Action | | | |
|------|-----------------|-------|--------|--|--|--|
| Canonical Exact | Canonical | 1 | ✏ Edit | ▢ Duplicate | ✓ Rule Attributes | 🗑 Delete |
| Canonical Nickname | Canonical | 2 | ✏ Edit | ▢ Duplicate | ✓ Rule Attributes | 🗑 Delete |
| Potential Distance | Potential | 1 | ✏ Edit | ▢ Duplicate | ✓ Rule Attributes | 🗑 Delete |
| Potential Nickname | Potential | 2 | ✏ Edit | ▢ Duplicate | ✓ Rule Attributes | 🗑 Delete |

1 of 1

# Set Rule Metadata

## Matchgrid: demo

COmanage™

**Manage**

**Display**

**Reconcile**

**Configure**

# Edit Canonical Exact

| Name * | Canonical Exact |
|---|---|
| Description | |
| Confidence Mode * | Canonical ∨ |
| Order | 1 ⌄⌃ |

*\* Denotes Required Field*

**SAVE**

# Define the search criteria for each attribute

## Matchgrid: demo

COmanage™

- Manage
- Display
- Reconcile
- Configure

COmanage Match > demo > Rules > Canonical Exact > Configure > Rule Attributes

# Rule Attributes

⊕ Add New Rule Attribute

| Attribute ▲ | Crosscheck Attribute | Search Type | Action | |
|---|---|---|---|---|
| dob | | Exact | ✏ Edit | 🗑 Delete |
| firstname | | Exact | ✏ Edit | 🗑 Delete |
| lastname | | Exact | ✏ Edit | 🗑 Delete |
| ssn | | Exact | ✏ Edit | 🗑 Delete |

1 of 1

# Searching for matches

Four search types are supported in v1.0.0 of COmanage Match:

- Exact
- Distance
- Mapping
- Substring

The search is partially configured in the attribute, and partially in the rule.

# Anatomy of an attribute

**Descriptive Information**

- Attribute Name
- Description
- API Name
- Attribute Group
- Display field in the Matchgrid?

**Matching Traits**

- General
  - Search Exact
  - Alphanumeric
  - Case Sensitive
  - Null Equivalents
- Search Distance
- Attribute Map
- Search Substring from/for

**Now these are important**

# Searching for matches
# EXACT

## Attribute Settings

| | |
|---|---|
| **Alphanumeric** | ☐ |
| **Case Sensitive** | ☐ |
| **Null Equivalents** | ☐ |

**For this attribute…**
- Should non-alphanumeric characters be stripped out?
- Does case matter?
- Is a blank "" the same as null 0000-00-00?

## Rule Setting

| | |
|---|---|
| **Attribute** * | firstname |
| **Crosscheck Attribute** | |
| **Search Type** * | Exact |
| **Required** | ☑ |

# Searching for matches
# MAPPING

## Attribute Settings

**Attribute Map**      Nicknames

**For this attribute…**
- What data dictionary (attribute map) should be used to find matches

## Rule Setting

**Attribute** *      firstname

**Crosscheck Attribute**

**Search Type** *      Mapping

**Required**      ☑

Searching for matches
# SUBSTRING

## Attribute Settings

| Search Substring From | 6 |
|---|---|
| Search Substring For | 4 |

**For this attribute…**
- Where in the field should the search start, and for how many characters - for example, search from 6 for 4 to compare values of the last 4 characters.

## Rule Setting

| Attribute * | ssn |
|---|---|
| Crosscheck Attribute | |
| Search Type * | Substring |
| Required | ☑ |

| | Rule Metadata |
|---|---|
| **Name** | `Canonical Exact` |
| **Description** | |
| **Confidence Mode** | `Canonical` |
| **Order** | 1 |
| **Attributes** | |

| | Attributes |
|---|---|
| **dob** | `Canonical Exact` |
| **Description** | |
| **Confidence Mode** | `Canonical` |
| **Order** | 1 |
| **Attributes** | |

|  | HR System of Record | Student System of Record |
|---|---|---|
| **Name** | Canonical Exact | Canonical Distance |
| **Description** | | |
| **Confidence Mode** | Canonical | Canonical |
| **Order** | 1 | 2 |
| **Attributes** | | |

# Testing

Bedrock of effective configuration

# Matchrid Records

# Resolving

When the humans get involved

# ...and now...

You know some.
What about the rest?

# This doesn't have to be the end of your COmanage Match journey!

**Technical Manual**

**User Guide**

…and other resources

- COmanage Mailing lists & slack
- Monthly Newsletter archive
- COmanage Open Office Hours (monthly)

- **Office Hours LIVE!**
  Wed 12:30 Governor's Square 14