# TIER ID Update Controller Architecture Flow with TIER ID Match

## *Refining the Registry Update Controller Process*

**DRAFT**

TIER Registry and DATA and API workgroups have provided an ID Match POC in September 2017.  This followed was work done in earlier projects and Benn Oshrin brought this out for demo purposes.  During December 2017 and January of 2018 refinement of a proposed architecture and of the POC ID Match has been underway.   Calling structures to the RESTFUL API and discussions regarding how it fits into the TIER Entity Registry - Identity OnBoarding process.  Availability of this is proposed ID Match component is projected to be during March 2018.

An architecture flow and event document for  TIER ENTITY REGISTRY - Identity OnBoarding  can be found at

TIER Enti...g (1).pdf

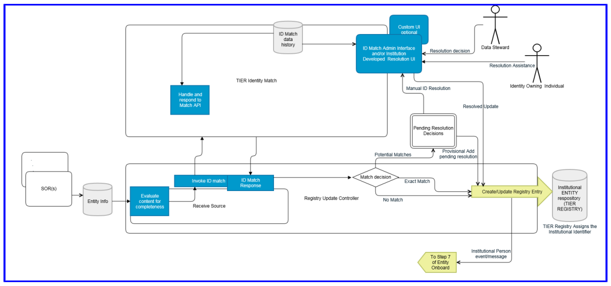
The intent of this document is to indicate in a high level sense how to align ID Match in your institutions processes.  It is intended to describe a pattern for institutions to use when building a Registry Onboarding process.

The ID Match is a restful API that can be called in many architectural patterns and or sources.   The Registry and API workgroups discussed patterns over several iterations and about 2.5 months have concluded with a suggestion as follows.   Calling ID Match component from the "Registry Update Controller" shown within the Identity OnBoarding flow.

## Considerations about ID Match

A simple manner to think about your identity system data flows can be considered in two fundamental flows, Identity Onboarding and Identity Provisioning and Authorization. Simply processes to get data into the ecosystem and processes to get data downstream or out to the applications relying on the access management services.

- Onboarding activity defines the entities in your Identity Service Ecosyystem.  Activities include population of you entity registry, assignment of an institutional identifier, matching input from multiple SORs, recording and /or assigning user accounts for the entity (Credential management Controller), triggering Subject info into the Groups Update Controller, and triggering the establishment of some  basis, and reference groups into the Groups component.  Onboarding does include the provisioning of downstream services.
- Provisioning and authorization provides support to set up Access management info and to distribute that downstream to the relying party applications. Builind application groups, extending and refining reference groups, setting rules in the grouping component.  Passing this information to the provisioning tools within the ecosystems and communicating those data to the services.
- Id Match Service is a portion of the onboarding process invoked by the Registry Update Controller.  The institution will need to make decisions on how to handle info returned by the ID Match Service.  Below multiple patterns will be described, each should be considered by your institution. Regardless of the pattern each institution should consider some basic questions. These are basic not necessarily simple.
    - Who will perform resolution activities resulting from ID match results?
    - What is the institutional support for reviewing the resolution activity for multiple possible  matches?
    - Is there existing or planned Governance for the Identity Service processes at your institution?
    - Is there Policy, Standards and procedure for Identity service Activities.
    - You will find other similar questions in setting the ID Match and Identity OnBoarding process.
    - The activity will likely need cooperation and input from management/staff managing the various SORs.
- Adding someone to the registry does not mean they have been given (provisioned access) .  rather it means they have been assigned an institutional identifier.  The grouping processes when forwarded into the provisioning space will create the access policy.


Custom UI optional
ID Match data history
ID Match Admin Interface and/or Institution Developed Resolution UI
Data Steward
Resolution decision
Resolution Assistance
Identity Owning Individual
TIER Identity Match
Handle and respond to Match API
Manual ID Resolution
Resolved Update
Pending Resolution Decisions
Potential Matches
Provisional Add pending resolution
SOR(s)
Entity Info
Invoke ID match
ID Match Response
Evaluate content for completeness
Receive Source
Match decision
Exact Match
No Match
Registry Update Controller
Create/Update Registry Entry
Institutional ENTITY respository (TIER REGISTRY)
TIER Registry Assigns the Institutional Identifier
Institutional Person event/message
To Step 7 of Entity Onboard

## Step-by-step guide

Determine your initial System of Record(s)  (SOR) to be onboarded or interfaced into your entity registry.

1. **Person is entered or updated (or self enters) into any of an  institutions SORs**.  The event/action needs to be captured and or acted upon for updates that must be forwarded to the registry.

    a. Map the attributes in your SOR into the normalized TIER information structure.  This would have been done during the process of configuring  the SOR to forward data to the registry.

        i. This structure will reflect the TIER minimal registry structure plus any additional person attributes or extensions you desire to store in your registry.
        ii. The TIER registry working group encourages you to maintain only data needed for access management in the registry.  OF course you may extend as needed of course.
        iii. It is intended that in Entities other than people will be accommodated by the registry.  Additional entities beyond person are intended to  be added in later work.  There are proposed structures for application entities in early stages of design.
        iv. Attributes are normalized into a consistent to allow the disparate data formats often encountered in each SOR to be brought into a single agreed upon form and format.  Normalized data is more easy to match and compare to locate existing individuals in the registry.

    b. Some institutions may have only a couple SORs while others could easily have 10 or more significant sources of persons being onboarded into a single Identity Registry.
    c. There is value in following the normalization and onboarding pattern even if your institution only has a single SOR at implementation time.  You will be ready and insured for the eventual addition of additional SOR(s).
    d. Activities above result in a connector to normalize and feed info from the Source to registry onboarding.

2. **Once you determine the data format.  Determine how you will trigger API call or messages** (choice is your institutions to make)

    a. When data changes in the SOR it must communicate that change to the registry.
    b. API or Messaging can be used to move the data.
    c. You must process the changed data (add, update and deletes), the lifecycles of individuals within the SOR and your institution.
    d. SIS, HR, Alumni, hospital staff HR, etc all may be managed in distinct manner by the specific SOR,  when moved to the registry the key is determining a unique person. Normalized data should be in place before matching is attempted.  The SOR itself is generally unaware of what (or if) another SORs may have already provided to the registry.
    e. Ideally the SOR systems move data in near real time to the registry.  If the SOR systems lag in time before sending you can expect to have issues in data currency and accuracy based on SOR data being consumed with time lags.  Depending on the specifics of your system cycles this is manageable, expected.

3. **Call the messaging or restful API to transmit data to the registry intake Receive Source Person Info**.
    a. This is well suited for a message based solution, TIER working groups have built demos with Rabbit MQ.  However,  you can choose other messaging tools or use an restful API approach if you prefer.
    b. Receive 'Source Person' info/message,  this component of the Registry Update Controller is responsible for
    c. Handling the SOR information as it arrives for processing.
        i. Is it correctly normalized
        ii. Is it "complete" to process further (and other such data handling checks
    d. Call the TIER ID Match component to see if this person has already been onboarded into the registry at a previous date time.
    e. step c.ican be deferred to step 4. but MUST be normalized prior to calling the ID Match.

4. **IdMatch  (Indentity Matching)**
    a. This is a restful API
    b. This component stores information related to previous matching history/experience.  These are identified by the idMatch reference identifier.
    c. Proposed matching attributes include: (taken from ID Match Attributes )
        i. Examples of common attributes for ID Match operations:
            1. Name
                a. Official
                b. Preferred
            2. Date of Birth
            3. Identifier
                a. National / SSN
                b. NetID / User ID
                c. Institutional Id (registry created)
                d. External Identifier (social id , eppn, ...)
            4. Email Address

- 5. System of Record label *("hrms", "sis", etc)*
- 6. System of Record identifier *(emplid, etc)*
- ii. Other types of attribute that can be used
  - 1. Street address
  - 2. postal code
  - 3. Telephone Number(s
- iii. IdMatch component allows configuration for attributes allowing for institutional custom for additional attributes.
- d. Id Match can return three successful responses. ID Match returns a result after recording the results internally.
  - i. This is a "new entity no match found",
  - ii. This entity matches this "existing registry entity matched exactly",
  - iii. This entity has "one or more potential matches" but Matching component needs help from a data steward.
  - iv. An ID Match reference identifier will be returned identifiers will be returned by the API for additional processing options.
  - v. ID Match can also return error state if improperly called.

5. **The onboarding Registry update controller will need to be configured to make DECISIONS and handle the results of IDMatch.**
   - a. New person no match found - the controller will invoke the no match to ADD a new person in the registry path
   - b. Existing registry person matches exactly - the controller will invoke the update the person in the registry path
   - c. Multiple possible matches requires additional insight of a person to do onboarding. Data Stewards should work with these entries to assure accurate result and minimize duplicates.
     - i. A choice to be made:
       - 1. Add the possible match in a "pending resolution" state to the registry. You may also want to mark all possible matches in a pending resolution state.
       - 2. Do not add to the registry. Wait for the data steward to review the data involved of the possible matches and the incoming data and then take action to move the data forward.
       - 3. Persons with appropriate knowledge at your institution can steward the data and provide the resolution in either case.
         - a. the pending resolution can allow for processes to move forward with the knowledge that this could be a duplicate and corrective action may be needed once the case is resolved.
         - b. do not allow pending entries to move forward. This avoids the need to undo processes that were allowed to proceed. It does create a delay in processes that consume the Institutional Person - pending resolution
       - 4. There are possibilities that the individual who is asking to be included in the identity system can respond to prompts to assist with the resolution. Wisconsin delivers the individual a "LINKING Key" to assist in resolving identity duplication. There have been other suggestions made to allow for the owner of the identity to participate in the resolution. It is necessary in this case to avoid disclosing info to the individuals. Design of this is very important so as to not allow a path for phishing or pharming of identity info. These are attempts to a decrease load on data stewards and to latency for resolutions.
       - 5. Institutions should carefully consider the choice of whether to use "pending resolution updates" or not. This is a key decision that has consequences regardless of the path chosen.
         - a. Delay entry to the registry until pending resolution is resolved
           - i. Agreed upon Data Stewards use an administrative User interface to review entities that had multiple possible matches.
           - ii. Staff need to have knowledge of the institution and follow agreed upon procedures that SOR and Institutional management have agreed upon.
           - iii. Institution needs to have a plan to provide resources for the activity.
             - 1. Identity Service and helpdesk staff.
             - 2. SOR staff from say registrar, HR etc
             - 3. An Identity services staff office empowered by agreed upon governance.
           - iv. Waiting for pending resolution can/will cause delay in individuals getting into the registry and assigned an institutional identifier, but this procedure will save on the downstream activities of individuals being propagated into downstream applications only to be deemed duplicate. Cleanup can be very expensive and time consuming.
           - v. cleaning up duplcate for the same person/entity can be can be an expensive backend process. This option front loads the decisions and greatly reduces the backend cleanup.
         - b. Allow entities that are pending resolution to update the registry before pending resolution is processed.
           - i. Entries are labeled with a pending resolution status.
           - ii. SOR information may indicate something about input quality let's call it "Questionable Providence"
             - 1. different SOR designations may imply a degree of trust related to the entity input from that SOR.
             - 2. registry entry does not directly imply access to services. Entities with pending resolution and /or questionable providence can be reflected into the groups update controller.
             - 3. These reference groups can then be used in access group policy for defining access to applications.
         - c. Hybrid approach that allows some Entities to enter the registry while others must wait for resolution processing to be added.
           - i. use a combination of the a and b method above.
           - ii. based on SOR and factors the institution may want to incorporate rules will be placed into the Decision processing in the Registry Update Controller.
           - iii. Example: Perhaps addition of a person attempting to take a DCE for payment activity is better off to be allowed to go forward to assure that the business transaction is completed and your institution gets credit/revenue for the course activity to be consumed. Blocking the addition in this case could drive your customer to another source of education and you lose opportunity. Allowing a pending resolution add to proceed is good since this creates revenue/service at a relatively low risk.
           - iv. Example: Someone joining in the HR area should likely be well identified before proceeding. Blocking is likely appropriate since cleanup could be expensive.
   - d. Manual Investigation of Pending resolution entries.
     - i. ID Match delivered Administrative User interface
     - ii. Institution developed User Interface
     - iii. Additional ID Match API Calls that can be useful.
     - iv. Institutions may decide to defer this processing for some administrative entries. See 5.c.i.4.b and 5.c.i.4.c options. Identities with status pending resolution.
   - e. Create / Update the registry Entry based on data steward decision

6. **Institutional Person established and change event communicated to remainder of identity onboard process. .**
    a. Institutional Person/Entity added event /message is communicated.
    b. see Identity Onboarding for steps 7 - 11 of onboarding (link at top of this page).

**Other considerations for managing the IdMatch service**.

1. Additional references from the TIER ID Match
    a. ID Match Attributes Review updated info for calling since the POC document.
    b. ID Match PoC  reviews the POC API calls and returns
    c. SOR-Registry Strawman ID Match API
    d. ID Match FLOW
    e. ID Match Scoping
    f. TIER Minimal Person Schema
    g. New person from SOR Narrative
2. Should the Institutional Identifier be contained in the ID Match data content with Institutional Identifier.
    a. This is backfed to the IdMatch data is a good feature for ongoing management between these data stores..
    b. This enables the ability to match on institutional id.  If SORs are aware of that id it provides a very reliable match.
    c. See the
3. Periodic updates of Id Match database from the registry.
    a. IdMatch  database contains certain information that should be periodically refreshed from the source.  (name, phone, etc)
        i. SOR-Registry Strawman ID Match API (see the Update Match attributes section)
        ii. periodic refresh
        iii. event based update when changes occur to the data elements in the Match database.
        iv. both ( this would be the preferred solution if it can be supported.)
4. UI for Identity Resolution
    a. Use default ADMIN UI in the TIER ID Match component.
    b. Develop a UI specific to your needsmore to come...

There are more steps (7 thru 11 from the identity onboarding flow architecture) to consider related to the Identity Onboarding, however since this document is focusing on the ID Match they are not covered here.  The Credential Management controller and the Groups Update Controller are treated in separate confluence documents.

## Related articles

- TIER Identity Onboarding - Credential Management Controller
- TIER ID Update Controller Architecture Flow with TIER ID Match
- TIER Update Controller Process with Id Match Flow Architecture