3.4 HRE Generic Object Data Model

Robin Lamacraft 2017-03-05

SCOPE

This describes how the higher-level user-entered data in HRE is stored in 2 basic types of structures. All records of the data have a type related persistent unique HRE-ID (and some other database management fields). The 2 types of structure are Nodes and Links. Each of these types is further specialized to satisfy the real world data model.

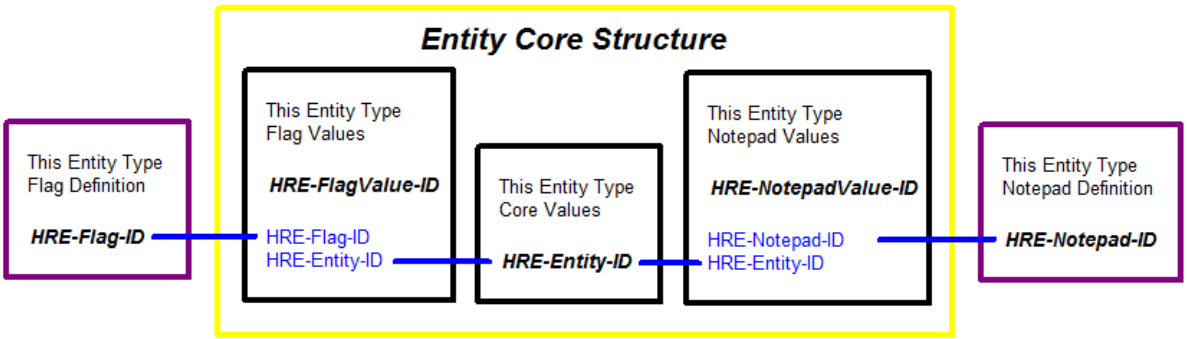
The end model has 10 user data Node Types.

- Group – Collections of objects that can be split or merged
- Item – Individual objects that can't be split or merged (e.g. artwork, image, etc)
- Person – Person (a special case of Item)
- Name – Name variants of a Node
- Location – Spatial data
- Source – Cited Evidence
- Repository – Where Evidence was obtained or located
- Event – a combination of activities in historical space and time involving many other entities based on evidence
- Task – a combination of activities in recent time, including research plans, progress and conclusions
- Occasion – a coordinate in historical time and space.

There is a restricted set of Links between these Node Types. There are a number of different uses for the Links between the Nodes. This document provides an overview of these differences. The aim is to design the Nodes and Links so that common functions only need to be described and implemented once.

Each Node may have its own collection of Name Variants.

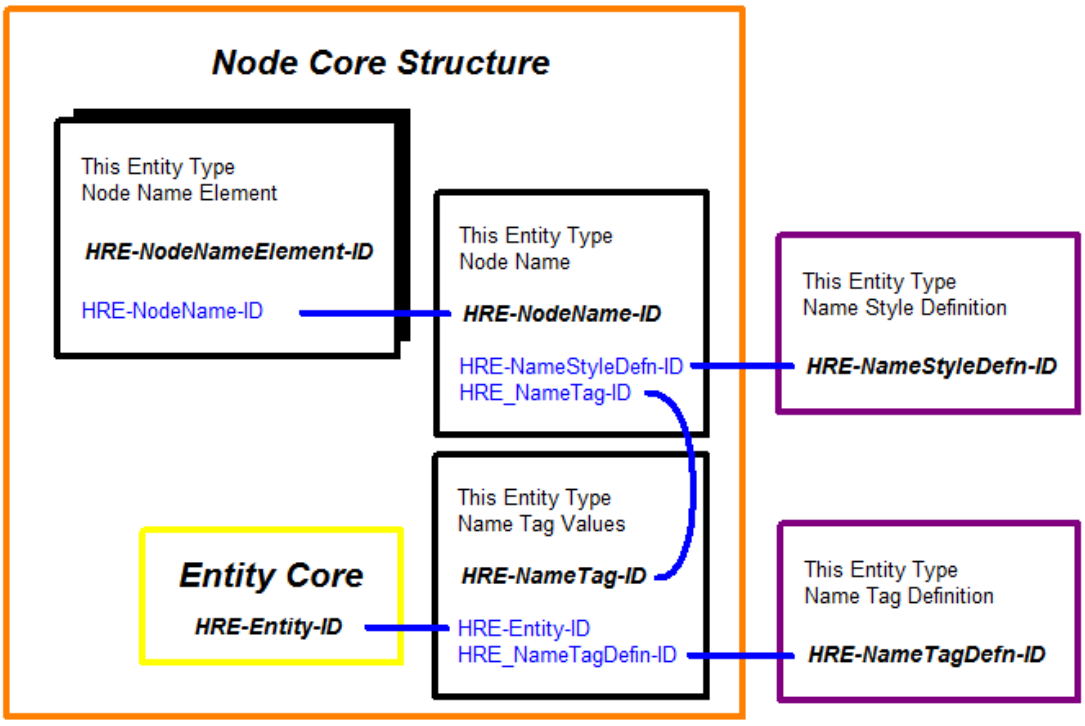
Each Node or Link may have its own collections of Flag and Notepad data fields.
[Flags are fields that have a limited number of alternative integer values which have text labels for each state – every instance record of a type has the same collection of flags.]
[Notepads are short strings that have no value constraints – except length. Notepads are stored HRE_ID indexed values – only non-blank values are stored.]



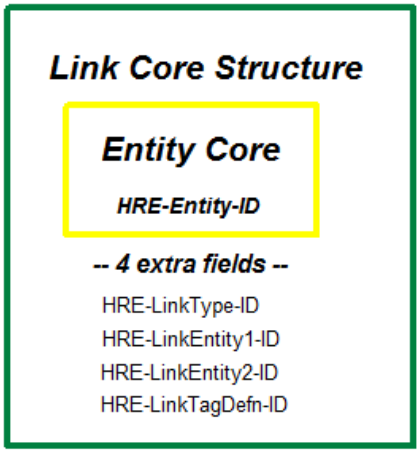
In the above diagram blue shadow denotes Flag and Notepad definition auxiliary data.
NOTE: Any description of the use of a text string below implies that that value can be an alternative of several translated values depending on various language selection settings.

Node Overview:
Nodes represent data that is free standing, like Persons, Sources, Locations and Events. They may have fields that refer to other objects (usually but not exclusively to Links). There are 2 classes of node; Item Nodes and Group Nodes:

- Item Nodes represent objects that can be split or merged like Persons, Artworks, etc. In this overview Images, External Files and Internal Texts are examples of Items.
- Group Nodes represent objects that can be split or merged like Collections, Events, Tasks, Locations, Sources, Repositories, etc
- Event, Task, Source and Repository Nodes have different linkage properties and management from all other Node Types. They act as major connection hubs
- Node Names - Many Node types have names. Some Node Names are controlled by Name Styles. Some Nodes can have a number of Names with qualifiers about their use and status. A Name Style can be as simple as a single text field to a collection of (say 20) text fields, some of which may have encoded data.



Link Overview:
These generally represent a pathway between 2 Nodes (but Citations are an exception – see later). A Link is characterized by having 2 identified fields that hold the HRE_ID of the object at each end of the Link.



- There are several types of Link:
    - Ownership Link: A directional link where one Node instance links to a Node of a type other than the owner's type. There may be many such links from the owner to similar Nodes. These can often be ordered, marked with priority of use. The nature of that linkage is under the control of a selected Tag Definition. An example of this is linking several names to a Person. In these cases the Person Name data shall only exist if the parent Person object continues to exist
      [A Tag Definition is a structure that controls the way the Link is controlled when it is created or edited and interpreted when output. A Tag Definition has Sentence Templates and Named Memos. The Tag definition may provide a different view of a link depending on which end of the Link is the focus of the reporting "X is father of Y" versus "Y is son of X". A Sentence Template is an encoded string that contains substitution variables to compose a compound narrative output including the content of the text of separately identified Named Memos]
    - Membership Link: A directional link where the objects at each end of the Link are both Group Nodes or the Group Node links to a Item Node as would a Sport Club (Group Node) link to a Player (Item Node). One Node instance links to an object of a type other than the owner's type
    - Relationship Link: Where one Node instance links to another instance of the same type, like Person to Person (father of) or Location to Location (near to). The nature of that linkage is under the control of a selected Tag Definition. In these cases the Node data can continue to exist if the Link is deleted
    - Associate Link: These only apply when one Node is an Event or a Task and the other Node is another type. The nature of that linkage is under the control of a selected Tag Definition
    - Citation: A Citation is a Link between a Source and any other type of Link or non-Source Node. It is not under the control of a Tag Definition. A Citation can be deleted without an impact on the Source or the other Node or Link.

NOTE: Each of these types of link will be discussed in the context of the links that connect to each of the 10 user data Node Types.

BETWEEN NODE TYPE LINKS AS TABLE

FROM NODE TYPE	TO NODE TYPE											
	Node Name	Group A	Group B	Location	Repository	Source	Event	Task	Person	Item A	Item B	Occasion
Group A	Owns	Builds	Groups	Groups	Groups	Groups	Groups Assoc	Groups Assoc	Groups	Groups	Groups	Groups
Group B	Owns	Groups	Builds	Groups	Groups	Groups	Groups Assoc	Groups Assoc	Groups	Groups	Groups	Groups
Location	Owns	In	In	Relates						Tag	Tag	Refs
Repository	Owns	In	In		Relates	Tag			(Refs)	Tag	Tag	(Refs)
Source	Owns	In	In		Tag	Relates			(Refs)	Tag	Tag	(Refs)
Event	Owns	In Assoc	In Assoc				Relates	Tag	Assoc	Assoc	Assoc	Owns
Task	Owns	In Assoc	In Assoc				Tag	Relates	Assoc	Assoc	Assoc	Owns
Person	Owns	In	In		(Refs)	(Refs)	Assoc	Assoc	Relates	Tag	Tag	
Item A	Owns	In	In	Tag	Tag	Tag	Assoc	Assoc	Tag	Relates	Tag	
Item B	Owns	In	In	Tag	Tag	Tag	Assoc	Assoc	Tag	Tag	Relates	
Occasion	Owns	In	In	Refs	(Refs)	(Refs	Owned	Owned				Relates

See next page for notes on these entries.

NOTES ON NODE TO NODE LINK TABLE:

1. "Owns" means that the TO Node is the owner of the From Node types
2. "Owned" means that the TO Node is owned by the From Node types
3. All Node Type can have Name Nodes. The naming link is under the control of a selected Name Tag Definition
4. Group A and Group B are 2 separate generic Group patterned objects (2 entries so that Group to Group Links can be shown)
5. "Builds" means that the groups are formed by grouping groups of the same type. The link is under the control of a selected Grouping Tag Definition
6. "Groups" means forming a group from objects that are not of the same as the parent group. The link is under the control of a selected Group Tag Definition
7. "In" means the Node is in a Group of that type. The link is under the control of a selected Group Tag Definition
8. Item A and Item B are 2 separate generic Item patterned objects (2 entries so that Item to Item Links can be shown)
9. The evidence types Images, External Files and Internal Text are examples of Items
10. "Relates" means that there is an asymmetrical links between 2 members of the same Node type (e.g. "father of" versus "son of "). This is a distinct type of "Tag" links where the nature of the link is under the control of a selected Relationship Tag Definition
11. "Assoc" means a non Event or Task Node type has a link to a Event or Task Node. The link is under the control of a selected Associate Tag Definition
12. "Refs" means that the TO Node <u>always</u> contains a reference to the FROM Node
13. "(Refs)" means that the TO Node <u>may</u> contain a reference to the FROM Node
14. "Tag" means a Link between 2 Node types is controlled by a selected Tag Definition. This includes a link between an Event Node and a Task Node

BETWEEN NODE TYPE LINKS

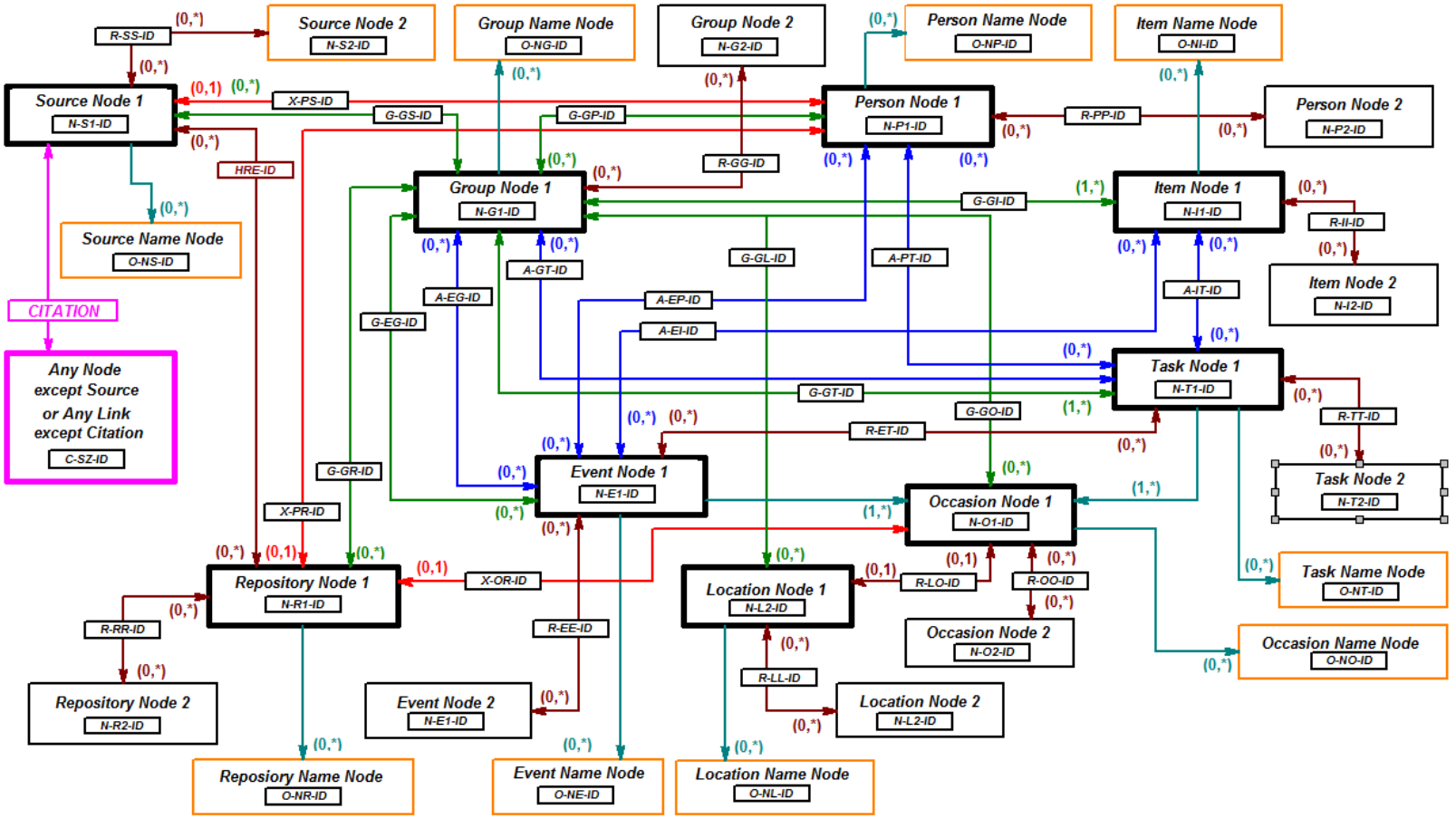


TABLE OF LINKS BY TYPE

- **OWNED (O) –** Name Variants for a Node Type, owned by that node instance. These OWNED links have a slightly different internal structure than other links:
    - **O-EN-ID: Event Name**
    - **O-IN-ID: Item Name**
    - **O-GN-ID: Group Name**
    - **O-LN-ID: Location Name**
    - **O-LO-ID: Occasion Name**
    - **O-NP-ID: Person Name**
    - **O-NR-ID: Repository Name**
    - **O-NS-ID: Source Name**
    - **O-NT-ID: Task Name**

- **GROUPED (G)** – Holds a membership of a Node instance of a nominated type to a Group of that Node Type:
    - **G-EG-ID: Event member of Group**
    - **G-IG-ID: Item member of Group**
    - **G-GG-ID: Group member of Group**
    - **G-LG-ID: Location member of Group**
    - **G-OG-ID: Occasion member of Group**
    - **G-PG-ID: Person member of Group**
    - **G-RG-ID: Repository member of Group**
    - **G-SG-ID: Source member of Group**
    - **G-TG-ID: Task member of Group**

- **ASSOCIATE** (A) – Holds data about the association of an instance of a Node Type in a role with an Event or Task:
    - **A-EG-ID: Group role in Event**
    - **A-EI-ID: Item role in Event**
    - **A-EP-ID: Person role in Event**
    - **A-GT-ID: Group role in Task**
    - **A-IT-ID: Item role in Task**
    - **A-PT-ID: Person role in Task**

- **EXTERNAL (X)** – Holds data about the association of an instance of a Location or Person Node Type to an instance of a Repository or Source:
    - **X-OR-ID: Occasion and Repository** - links to the location of the repository
    - **X-PR-ID: Person and Repository** - links the person who is the repository
    - **X-PS-ID: Person and Source**.- links the person who is the author or editor of a source

- **RELATIONSHIP (R)** – These link 2 instances of the same Node Type and stores data about the relationship. Relationships are often asymmetric. Hence the controlling information about that relationship type has separate relationship terms depending on which instance is the current focus.
    - **R-EE-ID: Event relationship to Event**
    - **R-ET-ID: Event relationship to Task**

    - **R-GG-ID: Group relationship to Group**
    - **R-II-ID: Item relationship to Item**
    - **R-LL-ID: Location relationship to Location**
    - **R-OO-ID: Repository relationship to Repository**
    - **R-SS-ID: Source relationship to Source**
    - **R-TT-ID: Task relationship to Task**

- **C-SZ-ID: CITATION (Z**) – These are links to an instance of a Source Node Type from:
    - An instance of any Node Type (but not a Source)
    - An instances of any Link Type (but not a Citation)

ADDING MORE ITEM AND GROUP NODE TYPES

When a new Node Type "W" is to be added to HRE the following other Nodes and Links need also to be added:

- The "W Definition" (Group or Item or Biological Item, etc)
- The "W" Node Type for the data of instances of "W"
- The "W-Name Style Definitions"
- The "W-Name Tag" Definitions"
- Any "W" data validation rules
- The Node Type "W Name" to hold the name data for "W" instances with Tag and Name Style
- The collection of tag definitions for "W" to W'' relationships
- The "W to W" Link to store relationships between "W" instances
- The collection of definitions for "W" instances to be Associates to Events
- The "W to Event" Link to store roles of "W" instances to Event instances
- The collection of tag definitions for "W" instances to be Associates to Tasks
- The "W to Task" Link to store roles of "W" instances to Tasks instances
- The collection of tag definitions for "W" instances to become members of a Group
- The "W to Group" Link to store roles of "W" instances in a Group

EXTRA ASPECTS IF ADDING A NEW GROUP TYPE

When a new Node Type "W" has Group properties then there needs to be extra Links to allow dor one Group Node Type to relate to a second or more Node Type(s) with Group properties.