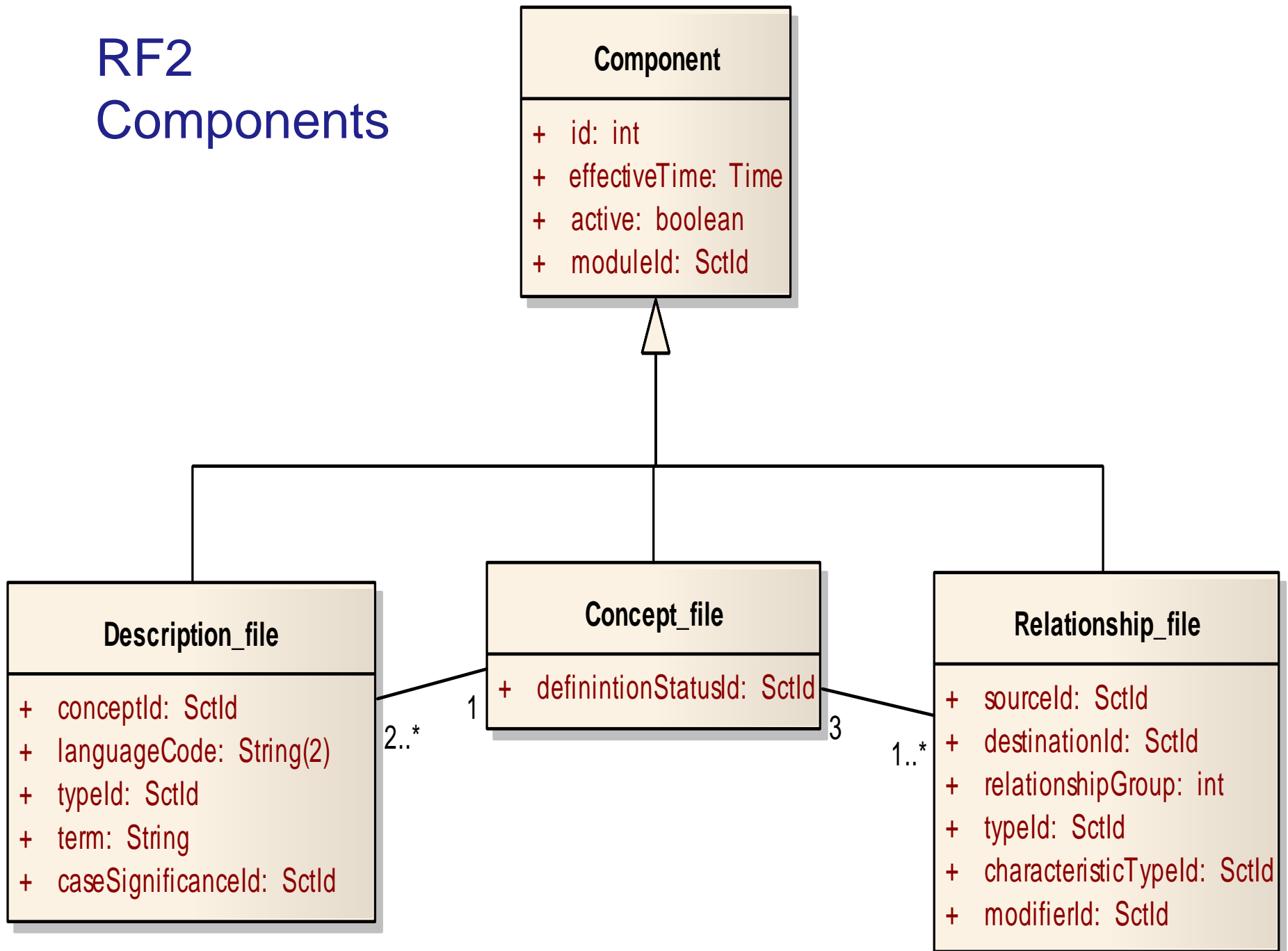


# Release Format 2 (RF2)

David Markwell  
IHTSDO - Chief Implementation & Innovation Officer

# RF2 Components





# Main changes in Release Formats 2 (1)

---

- Use of 'effectiveTime' to allow:
  - A single file/table to hold all released versions of SNOMED CT
  - Computation of snapshot view for any specified date
  - Delta release file of new rows supporting incremental updates
- Simple 'active' (boolean)
  - Replaces previous enumerated component status
  - Additional status information is provided by
    - Reference sets referring to reasons for inactivation
    - Reference sets that replace "historical relationships"



## RF2 - Component identification and versioning mechanism

---

Component
id: SctId effectiveTime: Time active: boolean moduleId: SctId

All components have

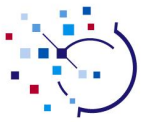
- **id**
  - Identifier of component
- **effectiveTime**
  - Time this version of the component superseded previous version
- **active**
  - Whether the component is for active use
- **moduleId**
  - Identifies extension version's module (supports development branching)



# Main changes in Release Formats 2 (2)

---

- Integer enumeration values replaced with Concept identifiers
  - Names of enumerated values are represented as descriptions associated with the identified concept. So these can be viewed and translated to other languages in the same way as other SNOMED CT concepts

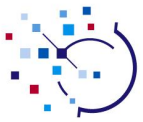


## RF2 – Replacement of enumerated values with meta-data concept identifiers

- Enumerated data type attributes in RF1 are replaced by SctId datatype in RF2
- These referring to a concept in the new 'meta-data hierarchy'
- These enumerations now have defined values that can be referenced and translated using standard SNOMED CT mechanisms

### Excerpt from top of new SNOMED hierarchy

- + **SNOMED CT Concept**
  - + <existing top-level hierarchy>
  - + **SNOMED CT Model Component**
    - + Core Meta-data concept
      - + Module
        - + IHTSDO maintained module  
SNOMED CT International
        - + UK NHS maintained module  
UK Drugs and Devices extension  
UK Admin extension
      - + Characteristic type (Relationship)
        - Defining
        - Qualifying
        - Additional
      - + Type (Description)
        - Fully specified name
        - Synonym



## Main changes in Release Formats 2 (3)

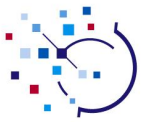
---

- **Description type**

- No longer indicates preference
- Preference
  - Is indicated by the relevant language Refset
  - Can be applied separately to each description type
- Extensible set using concept enumeration
  - Allows addition of new description types
  - New types can have longer term text (specified by metadata)

- **Language Code**

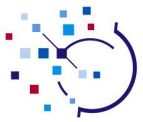
- Contains only language – not dialect
- Dialect is only represented in the Language Refset



## Main changes in Release Formats 2 (4)

---

- In RF1 'Subset' and 'CrossMapSet' tables provide specific metadata for subsets and cross maps
- In RF2 the 'Refset' is used as the single mechanism for referencing and adding information to SNOMED CT components
  - A Concept and associated Descriptions provide the name for the Refset
  - Hierarchical subtype relationships that determine the type of a Refset
  - Other metadata is provided by manifests represented as Reference sets



# Main changes in Release Formats 2 (5)

---

- Reference set members
  - Uniquely identified
    - By a UUID
  - Fully versioned using 'effectiveTime'
    - Can be inactivated using 'active' field value
  - Fully extensible with additional columns
    - Able to meet future requirements
    - Not limited by fixed specifications

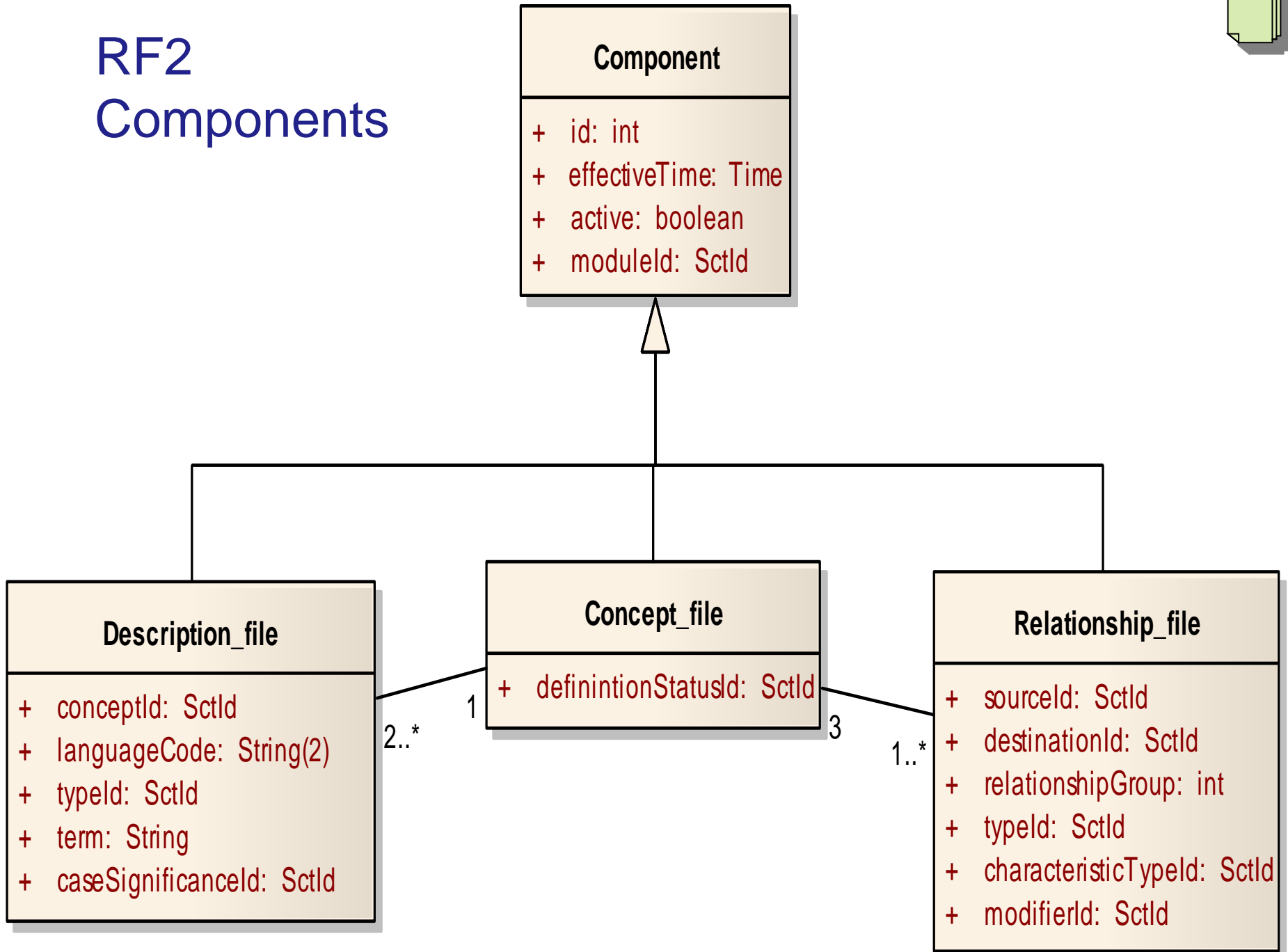


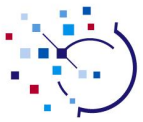
# Main changes in Release Formats 2 (6)

---

- Legacy identifiers Ctv3Id and SnomedId
  - Removed from the Concept file
- Mappings from concept.id to Ctv3Id and legacy SnomedId are represented using “Simple Map Reference sets”
  - One Refset for each of the legacy coding schemes

# RF2 Components

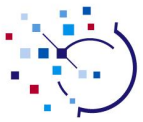




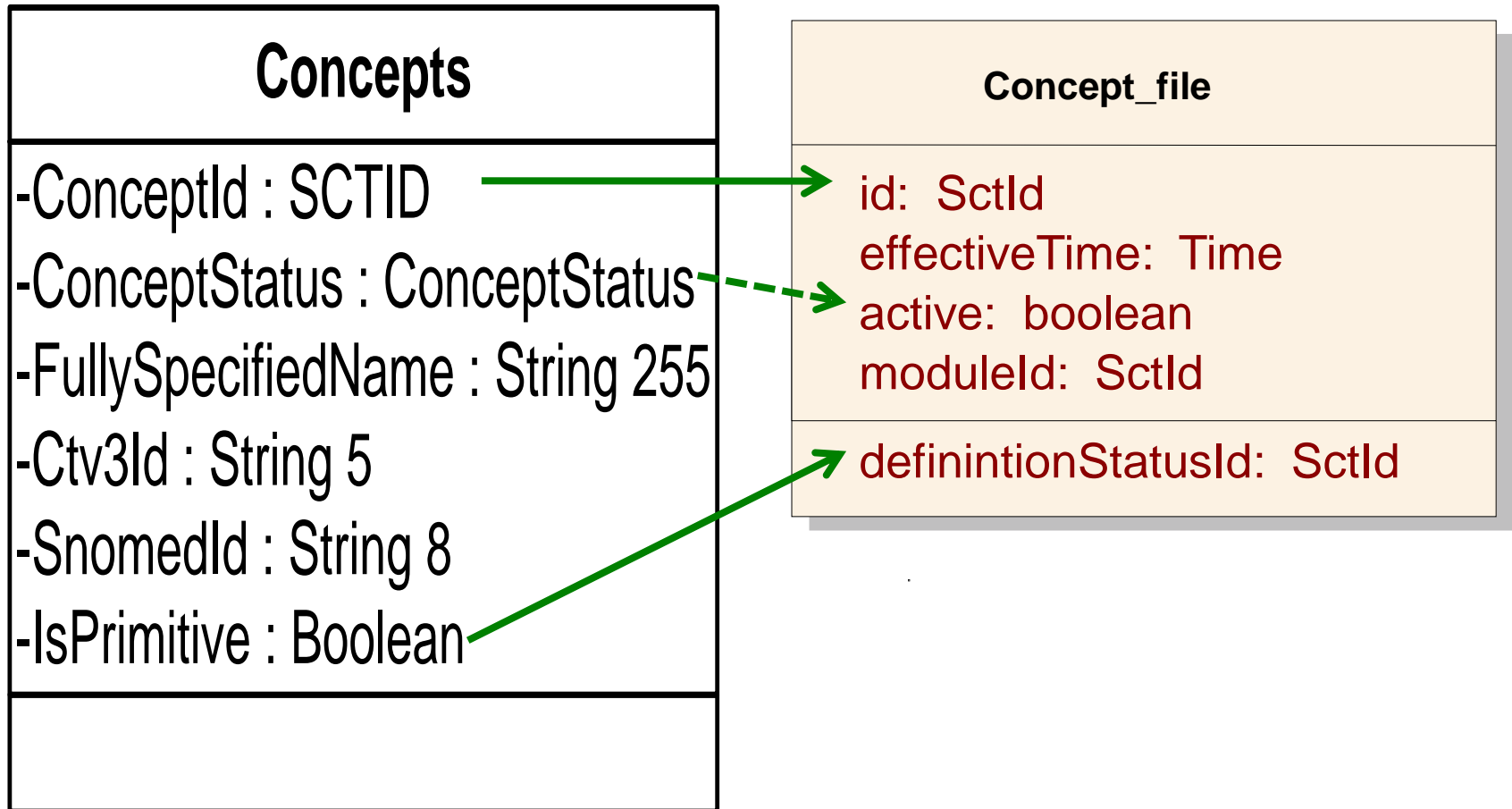
## RF2 – Other significant changes

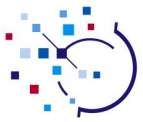
---

- **Separate Identifiers file**
  - For fixed unchangeable links between different identifiers of same component
    - E.g. between a developmental UID and other external identifiers
  - Note: Currently no rows in the released “identifiers file”.
- **Status representation**
  - ‘active’ attribute is boolean
  - Other status information (e.g. Ambiguous, duplicate, etc) is represented using Refsets
- **Historical relationships and references**
  - Replaced by Association Refsets
  - Not included in the relationships table



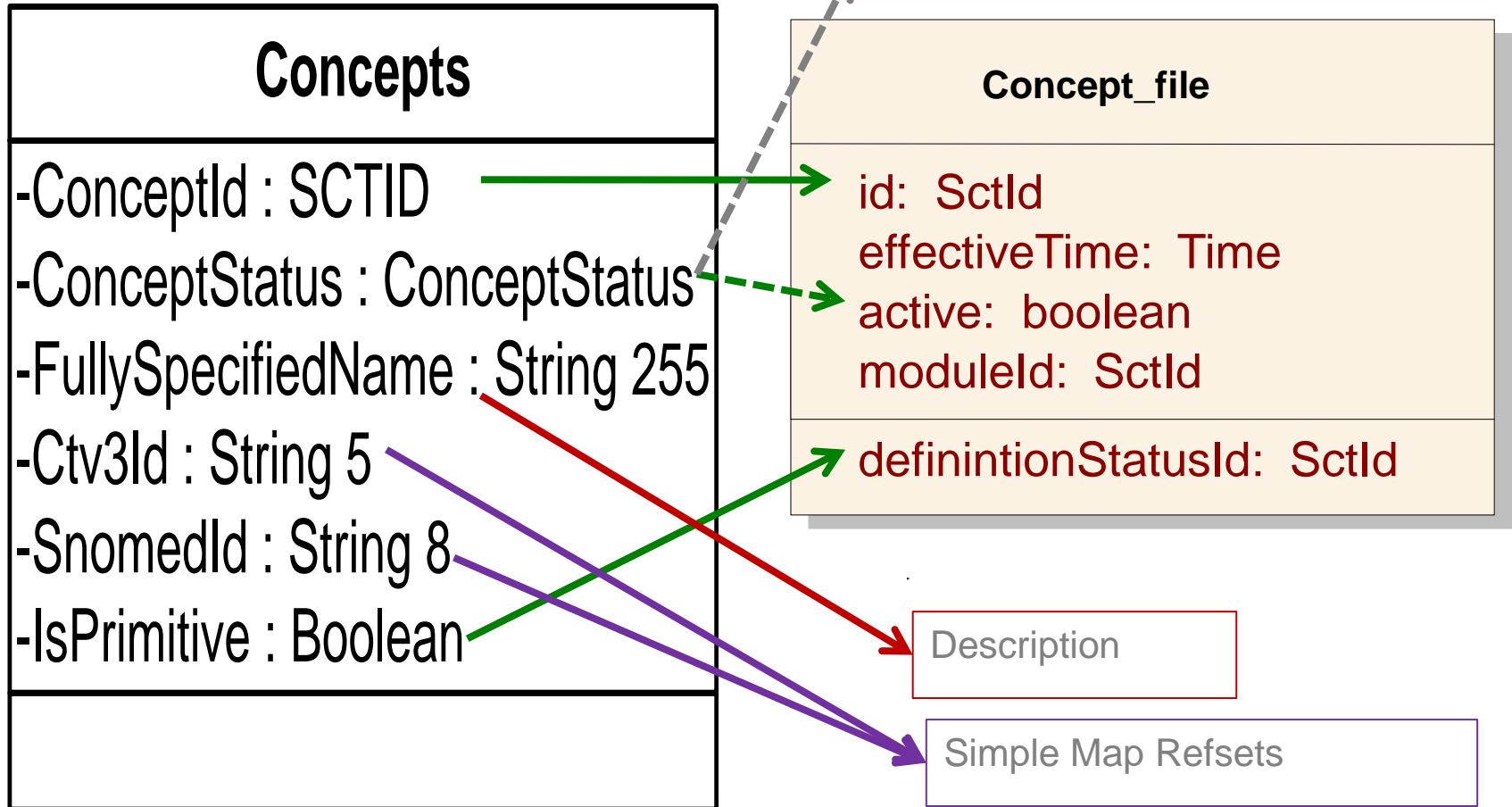
# Concepts RF1 → RF2





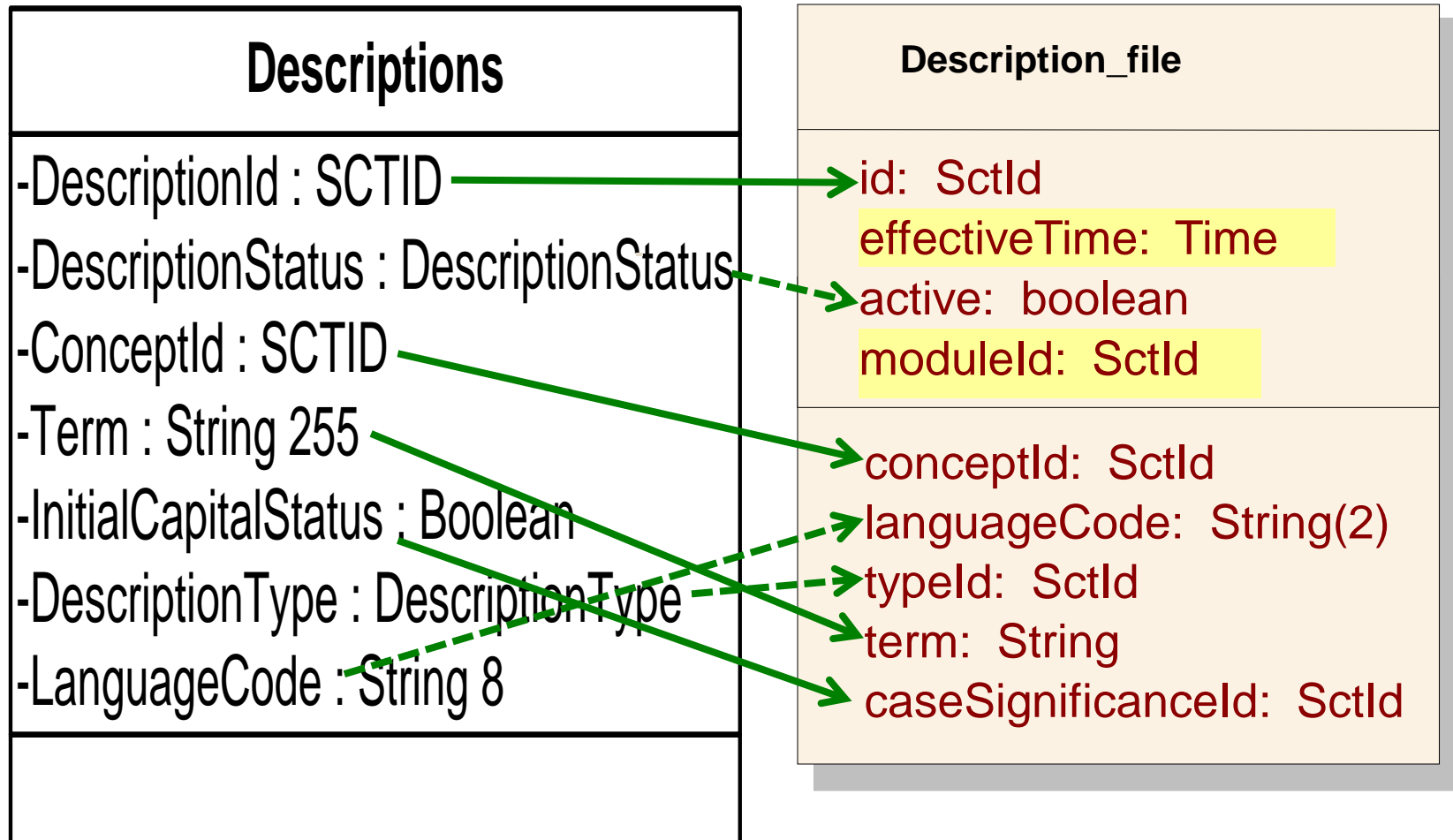
# Concepts RF1 → RF2

Concept Inactivation Refset



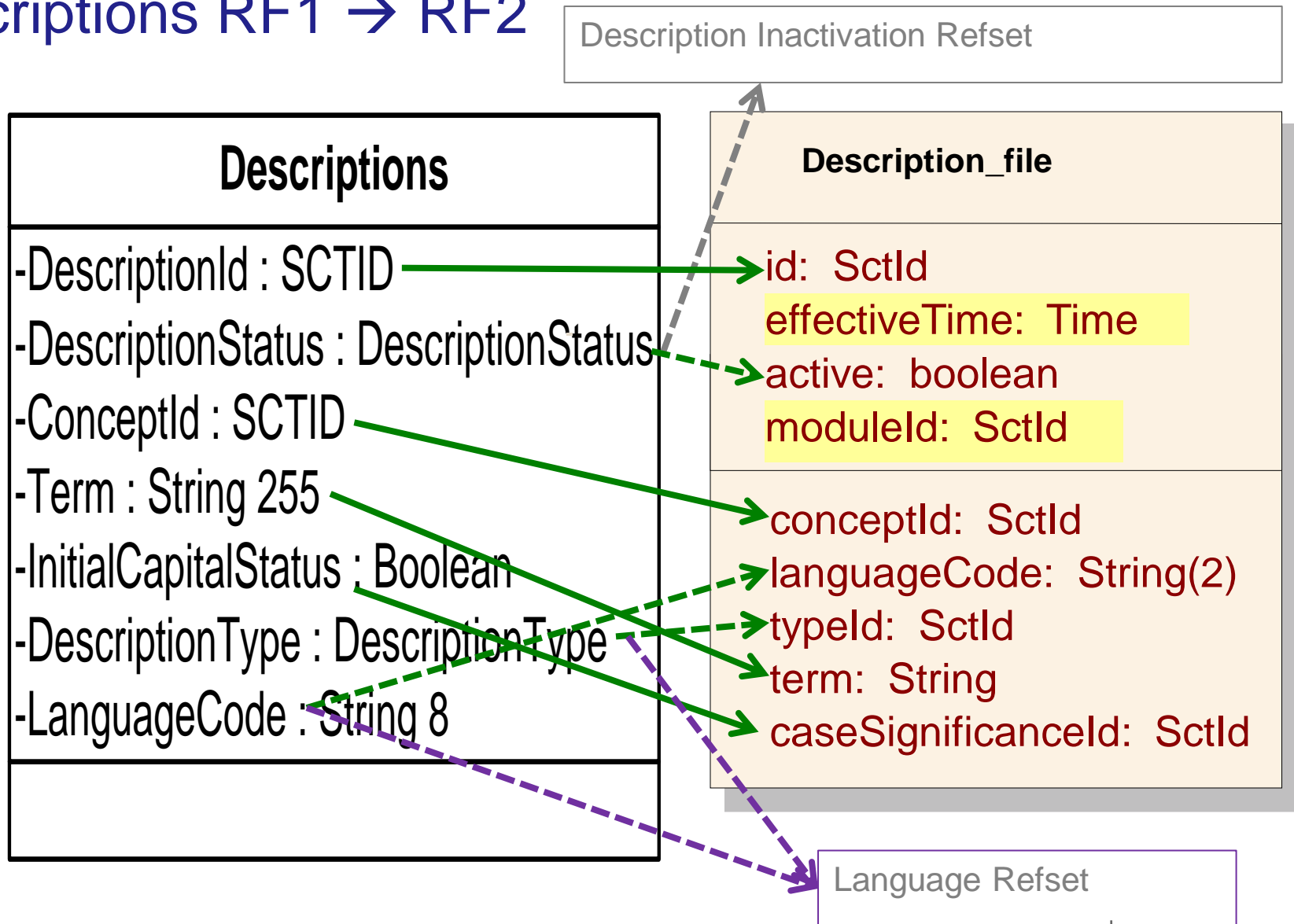


# Descriptions RF1 → RF2



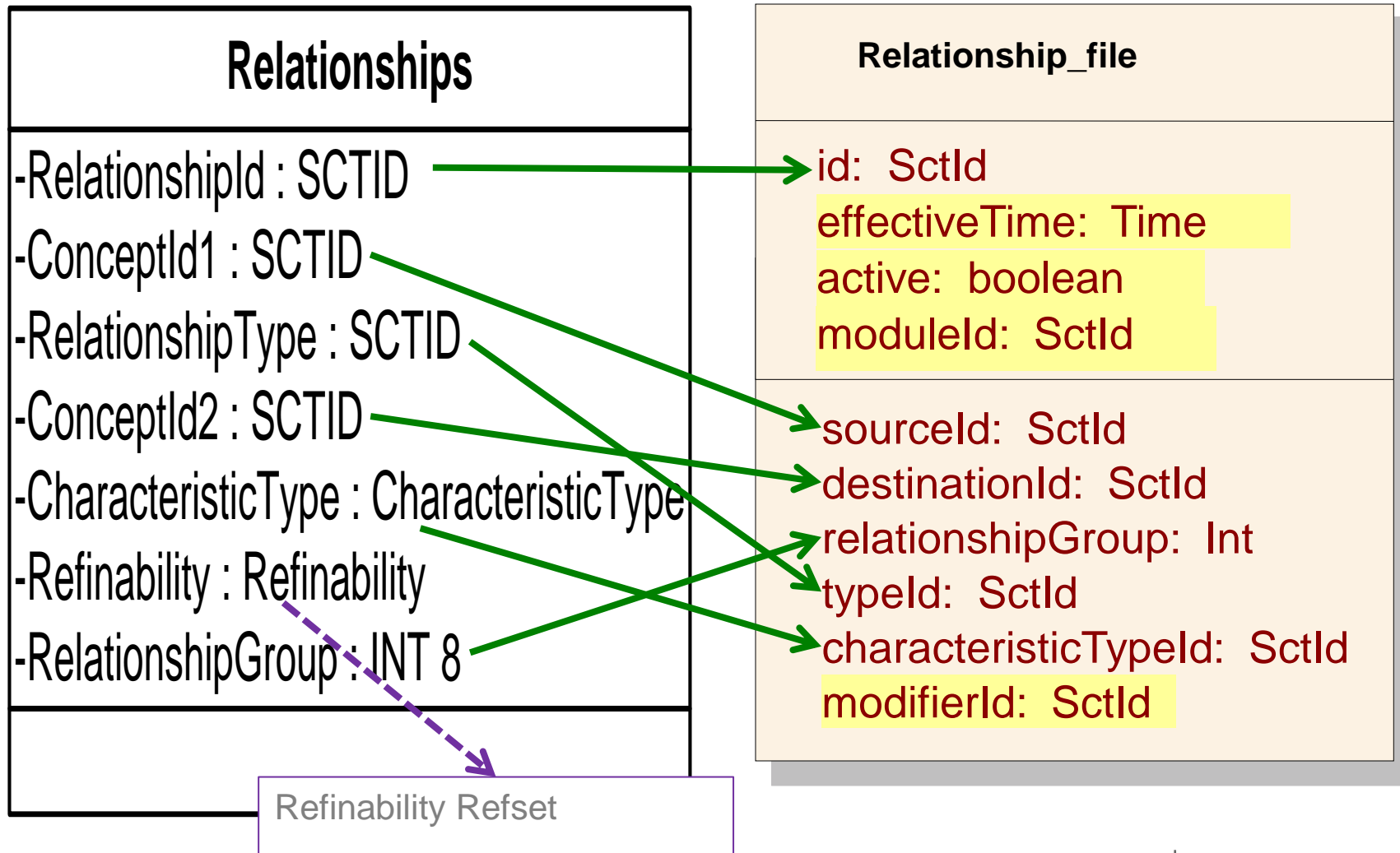


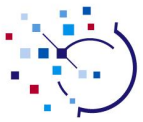
# Descriptions RF1 → RF2





# Relationships RF1 → RF2

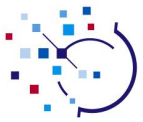




## Relationships RF1 → RF2

Relationships
-RelationshipId : SCTID
-ConceptId1 : SCTID
-RelationshipType : SCTID
-ConceptId2 : SCTID
-CharacteristicType : CharacteristicType
-Refinability : Refinability
-RelationshipGroup : INT 8

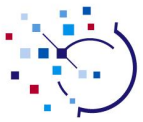
Relationship_file
id: SctId effectiveTime: Time active: boolean moduleId: SctId
sourceId: SctId destinationId: SctId relationshipGroup: Int typeId: SctId characteristicTypeId: SctId modifierId: SctId



# Benefits of RF2

---

- State-valid views of full release
  - Able to view any version of SNOMED CT by setting a date
  - Able to simply check changes made between any two versions
  - Common key and activity status structure on all components
    - id+effectiveTime+active
    - Reusable algorithm (e.g. SQL) process all component tables
- Reference Sets
  - General purpose extensibility approach
  - Common representational form for all derivatives
    - Subsets
    - Language and dialect configuration
    - Cross Maps
    - Alternative hierarchies
    - Annotations applied to concepts



# SNOMED CT Release Types

Release type	Description
Full	The files representing each type of component contain <b><u>every</u></b> version of every component ever released.
Snapshot	The files representing each type of component contain <b><u>one version</u></b> of every component released up to the time of the snapshot. The version of each component contained in a snapshot is the most recent version of that component at the time of the snapshot.
Delta	The files representing each type of component contain only <b><u>component versions created since the previous release.</u></b> Each component version in a delta release represents either a new component or a change to an existing component.

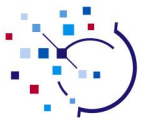
# Illustration of full release type

Components denoted by letters versions by digits



*Earlier* <-- FULL RELEASES FOR SEVEN RELEASE CYCLES --> *Later*

1	2	3	4	5	6	7
A,1	A,1	A,1	A,1	A,1	A,1	A,1
B,1	B,1	B,1	A,4	A,4	A,4	A,4
C,1	B,2	B,2	B,1	B,1	B,1	B,1
D,1	C,1	C,1	B,2	B,2	B,2	B,2
E,1	D,1	C,3	C,1	C,1	B,6	B,6
	E,1	D,1	C,3	C,3	C,1	C,1
	F,2	E,1	C,4	C,4	C,3	C,3
	G,2	F,2	D,1	D,1	C,4	C,4
		F,3	E,1	E,1	C,6	C,6
		G,2	F,2	F,2	D,1	D,1
		H,3	F,3	F,3	E,1	E,1
			G,2	G,2	F,2	F,2
			H,3	H,3	F,3	F,3
			I,4	I,4	G,2	G,2
				J,5	H,3	H,3
					H,6	H,6
					I,4	H,7
					J,5	I,4
						J,5
						K,7



# Reference Sets – What are they

---

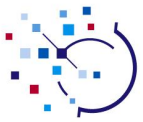
- Sets of references to SNOMED CT components
  - The name ‘Reference Set’ is abbreviated to ‘Refset’
  - Original SNOMED CT specifications used the name ‘Subset’ but this name only appropriate for some uses of Refsets
- Refsets are used for different purposes
  - To represent subsets
  - To indicate language/dialect preference for terms
  - To prioritize particular items in a search list
  - To specify alternative hierarchies
  - To attach metadata to a component
  - To attach annotations or other information to a component
  - To represent cross-maps
- Refsets are likely to have more uses in future



# Reference Sets – Common misunderstandings

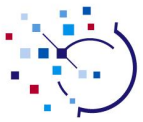
---

- **Refset** is not a new word for subset
  - **Refset** is a name for a flexible, extensible SNOMED CT file structure with many actual uses and many other potential uses
  - One use of a Refset is to represent a subset
  - Refsets do replace the RF1 SNOMED CT “Subset Mechanism” but that was badly named and Resets do even more “non-subset” stuff



# Simple Refset Structure

Field	Data type	Purpose
id	UUID	A 128 bit unsigned integer, uniquely identifying the reference set member.
effectiveTime	Time	Specifies the inclusive date at which this change becomes effective.
active	Boolean	Specifies whether the member's state was active or inactive from the nominal release date specified by the effectiveTime field.
moduleId	SCTID	Identifies the member version's module. Set to a child of [Module] within the metadata hierarchy.
refSetId	SCTID	Set to a child of [Simple type] in the metadata hierarchy.
referencedComponentId	SCTID	A reference to the SNOMED CT component to be included in the reference set.



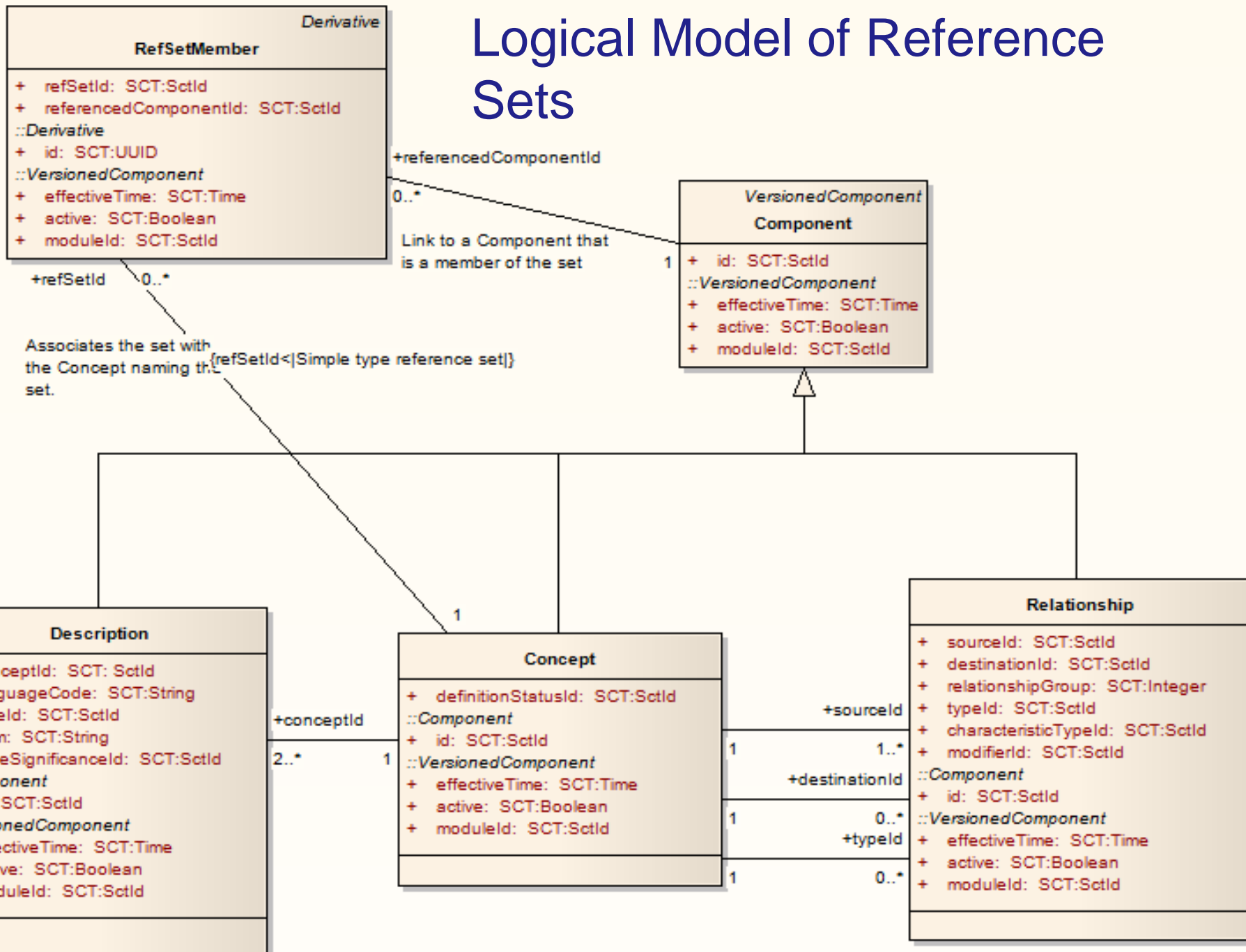
# Other Refset Structures

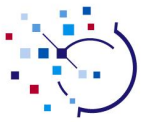
Field	Data type	Purpose
id	UUID	A 128 bit unsigned integer, uniquely identifying the reference set member.
effectiveTime	Time	Specifies the inclusive date at which this change becomes effective.
active	Boolean	Specifies whether the member's state was active or inactive from the nominal release date specified by the effectiveTime field.
moduleId	SCTID	Identifies the member version's module. Set to a child of [Module] within the metadata hierarchy.
refSetId	SCTID	Set to a child of [Simple type] in the metadata hierarchy.
referencedComponentId	SCTID	A reference to the SNOMED CT component to be included in the reference set.

- Plus sequence of one or more attributes each with is either an
  - SCTID;
  - Integer; or
  - String;(also used for times and UUIDs)



# Logical Model of Reference Sets

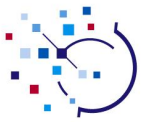




# Reference Set Types (1)

---

- **Simple Refsets**
  - Include/exclude
- **Language Refsets**
  - Support for description preference in languages and dialects
- **Ordered Refsets**
  - Prioritization of concepts or descriptions
  - Specification of a sort order
- **Hierarchical Refsets**
  - Alternative ordered navigation hierarchies
  - Aggregation hierarchies (e.g. for statistical classifications)
- **Cross Mapping**
  - Simple Map Refsets
  - Complex Map Refsets



## Reference Set Types (2)

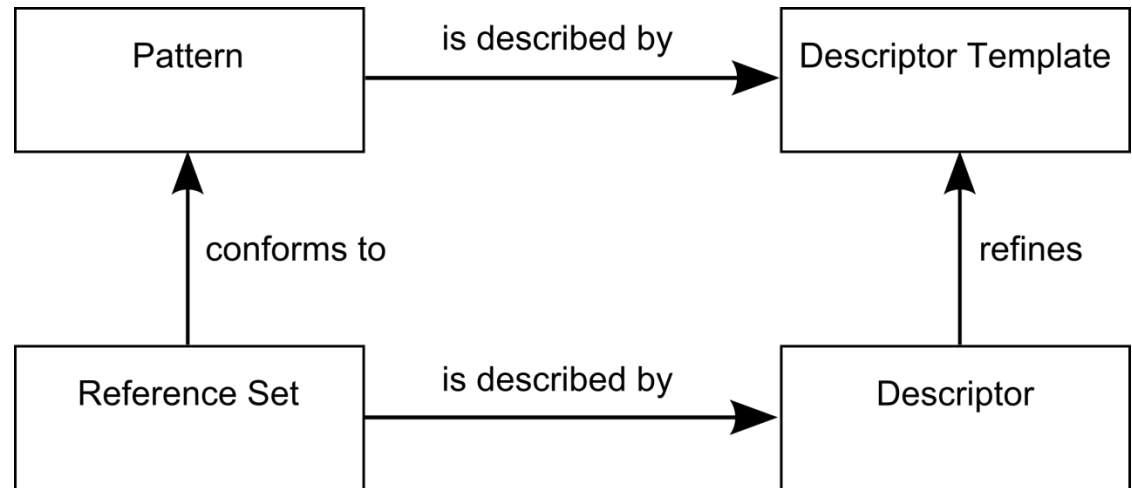
---

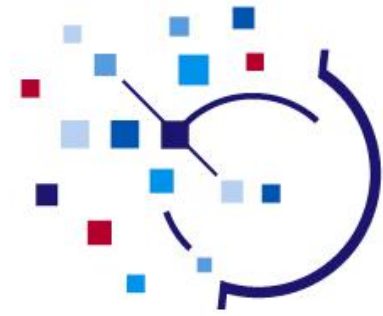
- **Component Inactivation Refsets**
  - Distinguish reasons for inactivation of concepts and descriptions (e.g. duplicate, ambiguous, etc)
- **Historical Association Refsets**
  - Refer to active components that represent inactive components (e.g. like RF1 Historical Relationships)
- **Query Refsets**
  - Represents and intensional subset/value-set
  - Syntax of query not yet specified by IHTSDO
- **Refset Descriptor Refset** (only one of these)
  - Specifies the field names and data types for a Refset (or for a set of similar Refsets)
- **Module Dependency Refset** (only one of these)
  - Indicates effectiveTime of other modules required to support content of a specific version of a module



## Refset Name and Structure Specification

- Refset Name is specified by a Concept with associated Descriptions
- Refset structure is specified by rows in the Refset Descriptor Refset





# Questions?