## Schema Data Format v3.0 Validation

Updated November 1, 2022

Document (top-level) objects:

* May contain *comment* and *privateData*.
* Error if any required elements are missing: *@id*, *sdfVersion*, and *version*.
* Error if *sdfVersion* does not equal "3.0".
* Error if both *entities* and *events* are empty or missing.
* Error if there are any duplicate 5-digit IDs within the *@id*s of all underlying *participants* objects.
* Error if there are any duplicate 5-digit IDs within the *@id*s of all underlying *relations* objects.

Event objects (*events*):

* May contain *children*, *children\_gate*, *comment*, *instanceOf*, *modality*, *outlinks*, *participants*, *privateData*, *relations*, and *repeatable*.
* Error if any required elements are missing: *@id*, *description*, and *name*.
* Error if there are any duplicate 5-digit IDs within the *@id*s of the array of *events* objects.
* Error if *children* is present but *children\_gate* is missing.
* Error if *children\_gate* (when present) does not match the controlled vocabulary: *and*, *or*, or *xor* (case-insensitive).
* Error if there is a mismatch between the number of *wd\_node*, *wd\_label*, and *wd\_description* values.
* Error if *modality* (when present) does not match the controlled vocabulary: *generic*, *hedged*, *irrealis*, or *negated*.
* Error if *instanceOf* does not match the *@id* of an *events* object.
* Error if both *children* and *wd\_node* are missing.
* Error if any element of *children* does not match the *@id* of an *events* object.
* Error if *outlinks* does not match the *@id* of an *events* object.
* Note: outlink cycles are NOT detected.

Participant objects (*participants*):

* May contain *comment*, *privateData*, *wd\_node*, *wd\_label*, and *wd\_description*.
* Error if any required elements are missing: *@id, entity*, *roleName*.
* Error if *entity* does not match the *@id* of an *entities* or *events* object.
* Error if there is a mismatch between the number of *wd\_node*, *wd\_label*, and *wd\_description* values.
* Error if two *participants* share the same values for both *roleName* and *entity*.

Entity objects (*entities*):

* May contain *comment* and *privateData*.
* Error if any required elements are missing: *@id*, *name*.
* Error if there are any duplicate 5-digit IDs within the *@id*s of the array of *entities* objects.
* Error if there is a mismatch between the number of *wd\_node*, *wd\_label*, and *wd\_description* values.
* Error if any required elements are missing: *wd\_node*, *wd\_label*, and *wd\_description.*

Relation objects (*relations*):

* May contain *comment*, *modality*, *name,* and *privateData*.
* Error if any required elements are missing: *@id*, *relationObject, relationSubject*, *wd\_node*, *wd\_label*, and *wd\_description*.
* Error if *wd\_node* contains more than one element.
* Error if there is a mismatch between the number of *wd\_node*, *wd\_label*, and *wd\_description* values.
* Error if *wd\_node* is from the temporal subset and *relationSubject* matches any *relationObject*.
* NOTE: cycles in temporal relations are NOT detected.
* Error if *relationObject* or *relationSubject* does not match an *@id* of an element in the *entities* or *events* arrays.
* Error if *modality* (when present) does not match the controlled vocabulary: *hedged* or *negated*.
* Warning if optional *name* element is missing.

Additional validation rules:

* Keywords of type *@id*: Error if value is not a [valid URI](http://sparql.org/iri-validator.html):
  + Includes *children*, *entity*, *@id*, *instanceOf*, *outlinks*, *wd\_node*, *relationObject*, and *relationSubject*.
* "**@id**": Error if @id is an invalid URI or does not follow naming convention (except for document IDs):

prefix:<KE\_type>/<unique-5-digit-number>/<anything>, where KE\_type is one of *Events*, *Entities*, *Relations*, and *Participants*.

* "**probability**": Error if value is not a float between 0 and 1.
* "**privateData**": Any valid JSON is allowed here. Outside *privateData*, unknown keywords (or known keywords in unexpected places) are rejected with an error.
* "**wd\_label/wd\_description**": Warning that empty values are ignored.