

{S}[B] SchemaBlocks

GA4GH Standards Documentation and Alignment Initiative

SchemaBlocks - History & Status

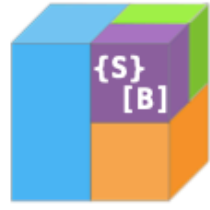


- Started by members of C/P & GKS, as *continuation* of former DWG Metadata work & other parts from GA4GH Schemas
 - core data model, objects
 - documentation
- Integration and exchange with *Phenopackets*, *Beacon* developments
- Maintained updated documentation and models in the Metadata repository
- December 2018:
 - first call with participants of different WS (GKS, C/P, Discovery)
 - launch of Github organisation “ga4gh-schemablocks”
 - New website @ schemablocks.org, with some initial documentation
- Early 2019: Alignment with Phenopackets & expanding engagement



GA4GH SchemaBlocks Home

SchemaBlocks is a “**cross-workstreams, cross-drivers**” initiative to document GA4GH object standards and prototypes, as well as common data formats and semantics.

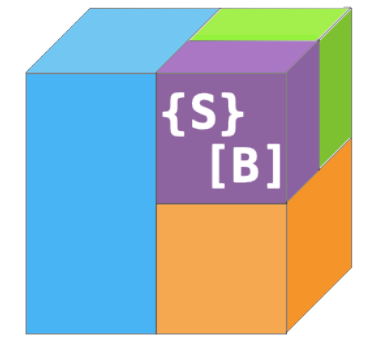


Launched in December 2018, this project is still to be considered a “community initiative”, with developing participation, leadership and governance structures. At its current stage, the documents can **not** be considered “**authoritative GA4GH recommendations**” but rather represent **documentation and implementation examples provided by GA4GH members.**

While future products and implementations may be completely based on *SchemaBlocks* components, this project does not attempt to develop a rigid, complete data schema but rather to provide the object vocabulary and semantics for a large range of developments.

The SchemaBlocks site can be accessed through the permanent link schemablocks.org. More information about the different products & formats can be found on the workstream sites. For reference, some of the original information about recommended formats and object hierarchies is kept in the [GA4GH Metadata repositories](#).

For more information on GA4GH, please visit the [GA4GH Website](#).



SchemaBlocks - A GA4GH Community Initiative

SchemaBlocks “Status” Levels

SchemaBlocks schemas (“blocks”) provide recommended blueprints for schema parts to be re-used for the development of code based “products” throughout the GA4GH ecosystem. We propose a labeling system for those schemas, to provide transparency about the level of support those schemas have from {S}[B] participants and observers.

Proposed {S}[B] Status Levels

The current status level of those recommendations is “proposed”.

- **playground**
 - early development or import stage, of any quality
 - no recommendation; existence does not mean any current or future {S}[B] support
- **proposed**
 - at least some {S}[B] contributors are in favour of such a block
 - the code may undergo considerable maturation
 - not recommended for integration into products w/o close tracking
 - contributions and discussions are encouraged
- **implemented**
 - mature block which is implemented in one or more {S}[B] aligned schemas
 - may be extended from a core block or be too specific for general (“core”) usability
- **core**
 - a schema block with recommended use
 - stable through minor version changes
 - has to be used in at least 2 standards/products approved by the GA4GH Steering Committee

SchemaBlocks^{{S}[B]} Mission Statement

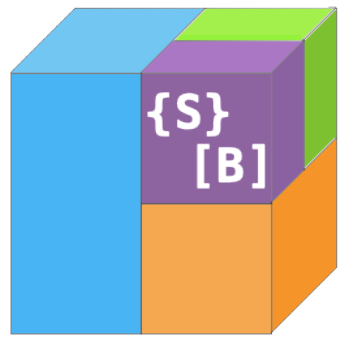
SchemaBlocks aims to translate the work of the workstreams into data models that:

- Are usable by other internal GA4GH deliverables, such as the Search API.
- Are usable by Driver Projects as an exchange format.
- Aid in aligning the work streams across GA4GH.
- Do not create a hindrance in development work by other work streams.

After discussions with stakeholders from GA4GH work streams and driver projects who create data models (such as Phenopackets, Search API) or who would use SchemaBlocks for the development of their APIs and data exchange formats (Beacon, EGA, GeL), the SchemaBlocks team has come up with the following principles for this initiative:

Work Stream Interactions

Work streams will continue to create standards proposals and their own coherent project implementations, but will work with the SchemaBlocks group to write the Blocks that will come from their own work and are considered of overarching use. Generally, primary work stream and driver project outputs will live in their own spaces outside of SchemaBlocks, with shareable, mature elements - code, documentation, implementation snapshots - being represented in {S}[B].



{S}[B] SchemaBlocks Github Repository Structure

- technical aspects of the SchemaBlocks initiative "live" in the {S}[B] Github organisation
- projects belonging to different classes
 - **blocks** projects for schemas & their conversion processes; general("blocks") and project specific ("sb-phenopackets")
 - **tools** project for schema conversion & validation scripts, configuration files
 - **ga4gh-schemablocks.github.io** website repository for "Markdown w/ YAML header" files for Jekyll based Github-Pages website generation
- currently, conversion of blocks source files to the different formats (JSON, Markdown, Markdown + YAML) is performed in a local setting (i.e. **tools** scripts are run across projects in a main {S}[B] directory)
- the main output (JSON, HTML) is accessible through **schemablocks.org**

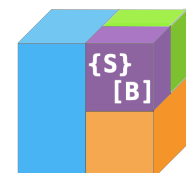
{S}[B] SchemaBlocks *JSON* *Schema* document format

- {S}[B] "blocks" are written in the YAML version of a JSON Schema document format
 - convenience choice - flexibility, readability, tooling ...
 - **not** implying specific semantics beyond some format conventions - extensible for use-case driven requirements
- the **meta** part (itself defined as a schema "block") contains housekeeping information
 - reference address & version
 - provenance & use cases
 - sb_status about "blessing level"
- the properties part defines the attributes including their description and usage examples
 - descriptions & examples provide the core documentation which is deparsed to the website documents

```
"$schema": http://json-schema.org/draft-07/schema#
"$id": https://schemablocks.org/schemas/ga4gh/AgeRange/v0.0.1
title: AgeRange
description: Age range
type: object
```

```
meta:
  contributors:
    - description: "Jules Jacobsen"
      id: "orcid:0000-0002-3265-15918"
    - description: "Peter Robinson"
      id: "orcid:0000-0002-0736-91998"
    - description: "Michael Baudis"
      id: "orcid:0000-0002-9903-4248"
    - description: "Isuru Liyanage"
      id: "orcid:0000-0002-4839-5158"
  provenance:
    - description: Phenopackets
      id: 'https://github.com/phenopackets/phenopacket-schema/blob/master/docs/age.rst'
  used_by:
    - description: Phenopackets
      id: 'https://github.com/phenopackets/phenopacket-schema/blob/master/docs/age.rst'
  sb_status: implemented
```

```
properties:
  start:
    allof:
      "$ref": https://schemablocks.org/schemas/ga4gh/v0.0.1/Age.json
      description: Age as ISO8601 string or OntologyClass
      examples:
        - age: 'P12Y'
    end:
      allof:
        "$ref": https://schemablocks.org/schemas/ga4gh/v0.0.1/Age.json
        description: Age as ISO8601 string or OntologyClass
        examples:
          - ageClass:
              id: 'HsapDv:0000086'
              label: 'adolescent stage'
          - age: 'P16Y6M'
  required:
    anyof:
      - start
      - end
  examples:
    - start:
        age: 'P12Y'
        ageClass:
          id: 'HsapDv:0000086'
          label: 'adolescent stage'
    end:
        age: 'P18Y'
```

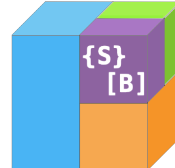
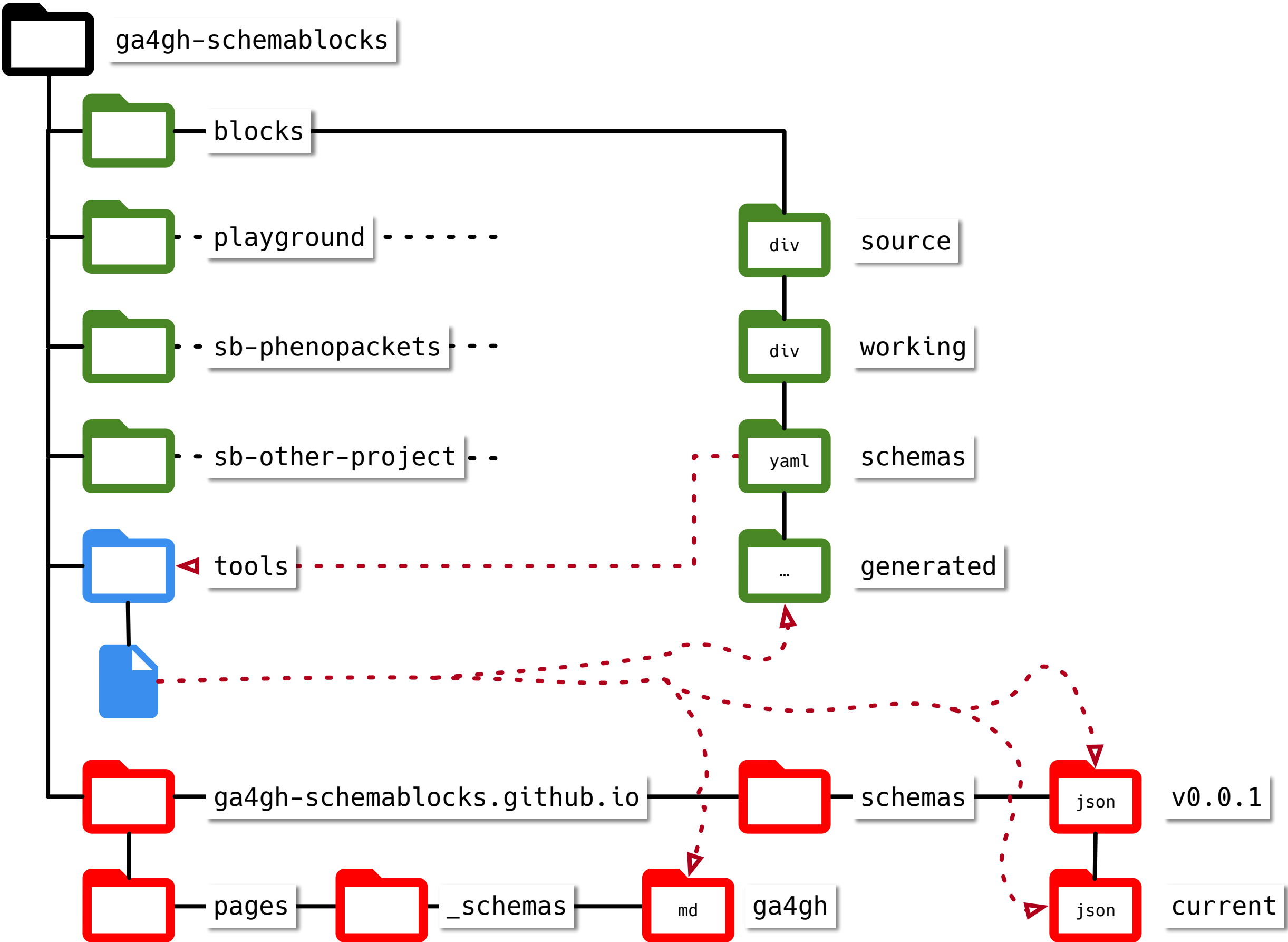


{S}[B] SchemaBlocks Github Repository Structure

blocks repositories

conversion/validation tools

website repository
(Markdown w/ YAML for Github Pages)



{S}[B] SchemaBlocks documentation generation

- block properties, their description and usage examples are represented on auto-generated web pages
- the (auto-generated) reference schema blocks themselves are accessible through canonical addresses in JSON format
 - ➔ <https://schemablocks.org/schemas/ga4gh/current/AgeRange.json>
 - ➔ used for accessing schema; validation ...

```
{
  "$id" : "https://schemablocks.org/schemas/ga4gh/AgeRange/v0.0.1",
  "$schema" : "http://json-schema.org/draft-07/schema#",
  "description" : "Age range",
  "properties" : {
    "end" : {
      "allof" : {
        "$ref" : "https://schemablocks.org/schemas/ga4gh/v0.0.1/Age.json",
        "description" : "Age as ISO8601 string or OntologyClass",
        "examples" : [
          {
            "ageClass" : {
              "id" : "HsapDv:0000086",
              "label" : "adolescent stage"
            }
          }
        ]
      }
    },
    "start" : {
      "allof" : {
        "$ref" : "https://schemablocks.org/schemas/ga4gh/v0.0.1/Age.json",
        "description" : "Age as ISO8601 string or OntologyClass",
        "examples" : [
          {
            "ageClass" : {
              "id" : "HsapDv:0000086",
              "label" : "adolescent stage"
            }
          }
        ]
      }
    }
  }
}
```

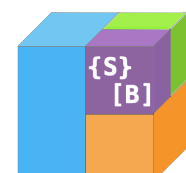
The screenshot displays the GA4GH SchemaBlocks website for the AgeRange schema. The page is organized into several sections:

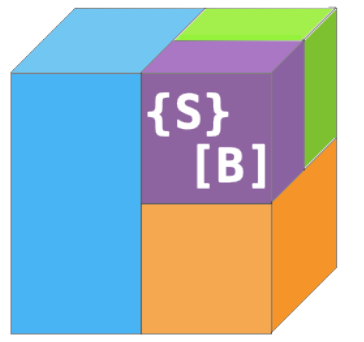
- GA4GH :: SchemaBlocks**: An Initiative by Members of the Global Alliance for Genomics and Health.
- Navigation**: News, Participants, Standards, Schemas, Examples, Guides & FAQ, Meeting minutes, Contacts.
- Related Sites**: GA4GH, GA4GH::Discovery, ELIXIR Beacon, Phenopackets, GA4GH::CLP, GA4GH::GKS, Beacon+.
- Github Projects**: SchemaBlocks, ELIXIR Beacon.
- Tags**: A grid of tags including Beacon, CP, Discovery, FAQ, GA4GH, GKS, MME, admins, code, contacts, contributors, coordinates, dates, developers, documentation, identifiers, issues, leads, news, phenopackets, press, tools, website.
- AgeRange**:
 - Status**: Implemented.
 - Provenance**: Phenopackets.
 - Used by**: Phenopackets.
 - Contributors**: Jules Jacobsen, Peter Robinson, Michael Baudis, Isuru Liyanage.
 - Source (v0.0.1)**: raw source [JSON], Github.
 - Attributes**: Type: object, Description: Age range.
 - Properties**:

Property	Type
end	https://schemablocks.org/schemas/ga4gh/v0.0.1/Age.json [SRC] [HTML]
start	https://schemablocks.org/schemas/ga4gh/v0.0.1/Age.json [SRC] [HTML]
 - end**: type: <https://schemablocks.org/schemas/ga4gh/v0.0.1/Age.json> [SRC] [HTML]. Description: Age as ISO8601 string or OntologyClass.
 - end Value Examples**:

```
{
  "ageClass" : {
    "id" : "HsapDv:0000086",
    "label" : "adolescent stage"
  }
}
```
 - start**: type: <https://schemablocks.org/schemas/ga4gh/v0.0.1/Age.json> [SRC] [HTML]. Description: Age as ISO8601 string or OntologyClass.
 - start Value Example**:

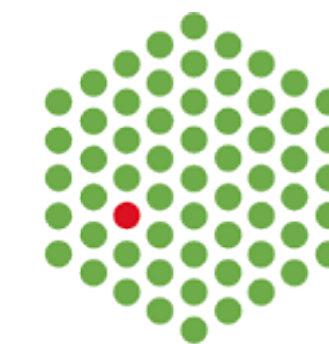
```
{
  "age" : "P12Y"
}
```



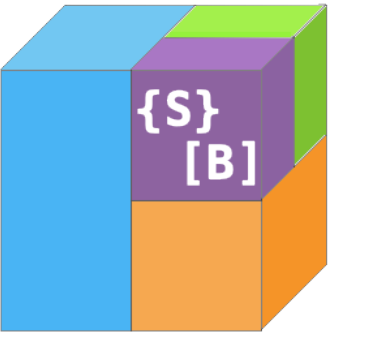


Future Development

- Inclusion into product approval processes?
- How to substantiate "GA4GH standard" claims?
- Lifecycle
 - Versioning and representation of donor schemas?
 - Development of conversion workflows for updated source products?
- Alternative/conflicting blocks...
 - Graded recommendations?
 - Name spacing?
- Management/Support?



{S}[B] Info



Leads

- Melanie Courtot [[↗](#)]
- Michael Baudis [[↗](#)]

Coordination

- Melissa Konopko

Websites

- schemablocks.org
- <https://github.com/ga4gh-schemablocks/>

Meeting minutes

- <https://schemablocks.org/categories/minutes.html>