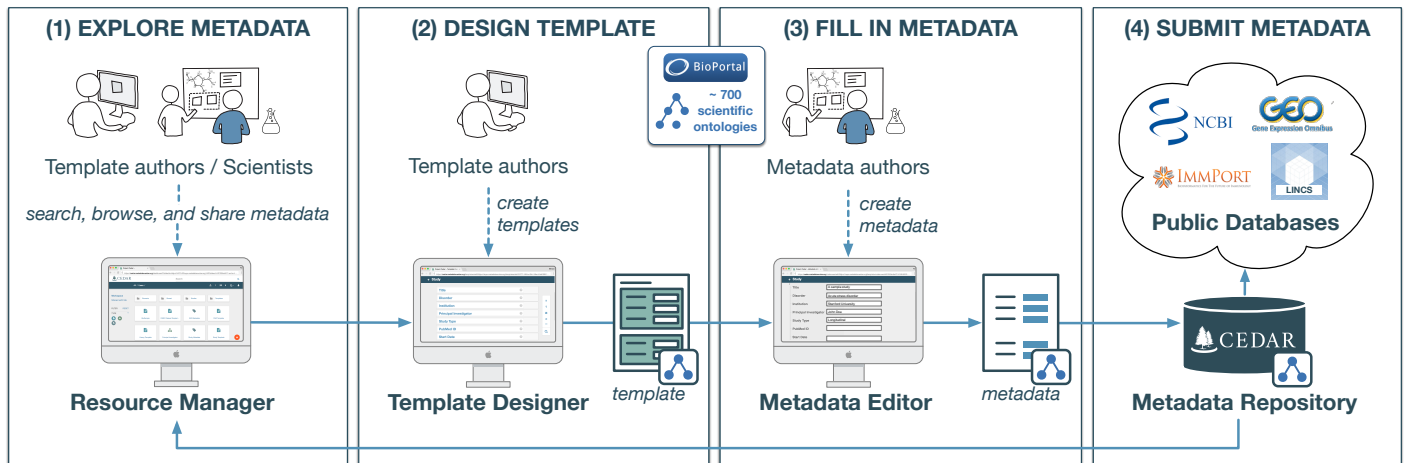# CEDAR: Web-Based Tools for Accelerating the Creation of Standardized Metadata

Mark A. Musen[1], John Graybeal[1], Martin J. O'Connor[1], Marcos Martínez-Romero[1], Attila L. Egyedi[1], Debra Willrett[1], S. Ahmad Chan Bukhari[2],
Steven H. Kleinstein[2], Kei-Hoi Cheung[2], Elizabeth Thomson[3], Alejandra Gonzalez-Beltran[4], Phillipe Rocca-Serra[4], and Susanna-Assunta Sansone[4]

[1]Stanford University; [2]Yale University; [3]Northrop Grumman, Inc.; [4]Oxford University

cedar.metadatacenter.org

CEDAR lets users create complete and comprehensive metadata for any scientific domain using standard ontologies and data structures.

(1) EXPLORE METADATA — Template authors / Scientists — *search, browse, and share metadata* — **Resource Manager**

(2) DESIGN TEMPLATE — Template authors — *create templates* — **Template Designer** — *template* — BioPortal ~ 700 scientific ontologies

(3) FILL IN METADATA — Metadata authors — *create metadata* — **Metadata Editor** — *metadata*

(4) SUBMIT METADATA — NCBI, GEO (Gene Expression Omnibus), ImmPort, LINCS — **Public Databases** — **Metadata Repository** (CEDAR)

WORKSPACE

Folders

CEDAR provides you a workspace to store your templates, elements and metadata. Organize your project into folders and manage your metadata over time..

BioSample Templates

Share

Collaborate with your team using groups. Share documents and folders in CEDAR with your team members. Publish your templates to the community by making them public.

Create a group

My Team

BioPortal

CEDAR is powered by BioPortal, the world's most comprehensives repository of biomedical and earth science ontologies, to provide controlled terms for metadata.

Ontology: NCBITAXON
- Entity
- Conceptual Entity
- Physical Object
  - Substance
  - Anatomical Structure
  - Manufactured Object
  - Organism

TEMPLATES

Templates

Building templates involves composing forms of fields, each with a name and description.

Enter Field Title
Sample Name

Enter Field Description
Enter the name of your sample.

Basic field types include short and long text, numbers, email, and more.

📅 Date		- - Break	
@ Email		▶ Video	
# Number		🖼 Image	
▦ Phone		⚎ Element	

Elements

Elements are reusable groups of fields and other elements. Elements can be shared on CEDAR and nested to any level. Using elements enables templates to be composed of community-defined building blocks to meet common specifications or needs.

- Study
  - Principal Investigator
    - First Name
    - Last Name
    - Address
  - Publication
    - Title
    - Journal
    - Date

METADATA

Metadata

CEDAR supports features that allow easy and rapid entry of metadata. Computer-assisted intelligent authoring provides suggestions for both fields and values.

BioSample Human

Tissue: brain
Sex
Isolate*
Age*
Biomaterial Provider*

Optional Attribute

disease

Value
- astrocytoma (DOID; 19%)
- Parkinson's disease (DOID; 19%)
- central nervous system lymphoma (DOID; 10%)
- oligodendroglioma (DOID; 9%)
- bipolar disorder (DOID; 9%)
- schizophrenia (DOID; 9%)
- neuroblastoma (DOID; 6%)

JSON-LD

Templates and metadata are represented as JSON-LD, a machine-readable and widely accepted interchange format on the Web.

"organism": {

"@type": "http://data.bioontology.org/ ontologies/NCBITAXON/classes/ http%3A%2F%2Fpurl.bioontology.org%2 Fontology2FSTY%2FT001",

"@id":"http://purl.obolibrary.org/ NCBITAXON/9606" }

VALIDATE and PUBLISH

Validation

Automatic validation is done on the fly to give immediate feedback of errors. The validator framework wraps other existing validation services, e.g., the BioSample Validator.

⚠ BioSample validation errors:
- Empty attribute value for attribute 'tissue'.
- Sample has missing mandatory attribute(s). If you do not have information for the required field(s), please provide value as either 'not collected', 'not applicable' or 'missing'.

REST APIs

Create, read, update and delete with CEDAR REST API microservices. Send metadata to your repository or app, or publish to CEDAR!

DELETE	/folders/{id}
GET	/folders/{id}
GET	/folders/{id}/details
POST	/folders

Adoption and Evaluation

BD2K - LINCS DATA COORDINATION AND INTEGRATION CENTER, AIRR Community MiAIRR, Stanford LIBRARIES, RDA RESEARCH DATA ALLIANCE, Open ImmPort, NIH Commons Credits Pilot, National Center for Biotechnology Information NCBI