

# Product Sheet for the Cancer Models Database (caMOD)

[Functional Capabilities](#)

[Tasks users seek to perform and types of problems addressed using the resource](#)

[Intended users](#)

[Animal models covered](#)

[Technical Data and Architecture](#)

## Functional Capabilities

- Brief description of function and scope  
*The cancer models database (caMOD) is a web-based resource that provides information about rodent models for human cancer to the public research community. caMOD provides the following key capabilities to its users:*
  - **Data Submission**--Data in caMOD are extracted from the public scientific literature by curators and verified by the scientists who generated or worked with the models, or they are directly submitted by scientists.
  - **Cancer Model Search**--Users can retrieve information about the making of models, their genetic descriptions, histopathology, derived cell lines, associated images, carcinogenic interventions, microarray data, and therapeutic trials in which the models were used. caMOD provides links to PubMed for associated publications and other resources such as mouse repositories, detailed information about altered gene, pathway affected, and information about human clinical trials that utilize the same compounds as the pre-clinical trials in the animal models.
  - **System Function Administration**--The Admin function provides services for user registration, review of submitted models and database management.  
caMOD is accessible at: <http://cancermodels.nci.nih.gov>.

## Tasks users seek to perform and types of problems addressed using the resource

Public data submission: Users can submit data about animal models for cancer. The database is equipped to store information about mouse, rat, zebrafish and can be adopted for other model species. The submission is divided in several categories which can be used or not depending on the way the model was generated.  
This table summarize the pages of caMOD that can be used to enter or review data about a model submitted into caMOD

Information Categories in caMOD	Description of Category's Focus
Model Characteristics	Provides an overview about the experimental design and phenotype of the model.
Genetic Description	Describes genetic modifications of the model e.g. transgene, targeted modification, heritable mutations induced by chemicals or radiation
Carcinogenic Exposure	Describes exposure to carcinogens e.g. chemicals, radiation, hormones that induce or accelerate tumor development. These changes are not inheritable.
Transient Interference	Describes morpholino and siRNA experiments
Xenograft / Transplant	Describes xenograft model or transplantation experiments
Publications	Lists publications written about the animal model and experiments conducted with the animal model.
Histopathology	Stores histopathology results for each affected organ in the model

Cell Lines	Describes Cell Lines that were generated from the model
Therapeutic Approach	Illustrates experiments and the results of attempts to cure the animals
Image and Microarray Data	Repositories for appropriate images and expression data for the model
Model Availability	Allows the user to indicate sources for acquiring the animal model itself or other strains that were used to generate the animal model

The model characteristics page provides an overview. The genetic description, carcinogenic exposure, transient interference, and xenograft/transplant categories describe how the model was generated. The user can make **multiple entries for each category**. For example, if the model was exposed to two different chemicals the submitter would enter data for each instance.

*Example:*

A double-transgenic animal was crossed with a knock-out animal, treated with UV-light. To properly describe this model, you would make the following entries:

- Genetic Description > Enter Engineered Transgene > Transgene 1
- Genetic Description > Enter Engineered Transgene > Transgene 2
- Genetic Description > Enter Targeted Modification > Knock-out Gene
- Carcinogenic Interventions > Enter Radiation > UV-light

Data Queries:

In caMOD, users do not need accounts to query the database. All searches can be performed without being logged in or even without a user account.

Search results are presented in a customizable list from which the detail pages for each model can be accessed. The search detail pages are organized by the same categories as the submission (see above). Users with user accounts can build and store queries.

Each search detail page offers users the opportunity to comment on the entry or add data to a particular record. The comments will be screened before they appear online (for more detail see Administrative features)

There are three different search modes for querying caMOD: simple search, advanced search and search via the table of contents. The drug screening search, a fourth search option, allows users to search for a specific compound or drug and results of its use in yeast screening, in-vivo screening, pre-clinical trials and in clinical trials.

*Simple Search*

The simple search page lets the user query by keyword, model name, principal investigator, site of lesion/tumor, and species.

*Advanced Search*

The advanced search page lets the user query for the following in addition to the criteria listed for the simple search:

- Diagnosis
- Phenotype
- Gene Name for modified genes
- Type of modification
- Genomic segment modifier
- Inducing agent for heritable induced mutations
- Carcinogens to which models were exposed
- Name of cell lines generated from the model
- Models with associated therapeutic approaches data
- Models with associated microarray data
- Models with associated transient interference data

- *Models with metastases*

#### *Table of Contents*

*The Table of Contents is organized according to animal model organ systems or other specific questions of interest e.g. animal models which develop metastases. The table of contents consists of predefined searches which can be activated by clicking a specific hyperlink.*

#### *Drug Screening Search*

*The Drug Screening Search page allows users to search for a specific compound or drug and results of its use in yeast screening, in-vivo screening, pre-clinical trials and in clinical trials. The yeast and in-vivo screening data were obtained from the NCI Developmental Therapeutics Program (<http://dtp.nci.nih.gov>).*

#### *Administrative features*

*Before model entries are displayed on the search pages they have to undergo a two-tier review process. caMOD Admin functions include responsibilities for three assignees:*

*Coordinator: The coordinator can add new users, appoints screeners and editors, assigns models to screeners and editors, and assigns comments to a screener.*

*Screener: The screener scans submitted, completed records for appropriateness. The screener can approve or reject a record. The screener scans comments for appropriateness. The screener can approve or reject a comment.*

*Editor: The editor checks the record for completeness and correctness. The editor approves the record for display or asks the submitter for more information.*

#### **Intended users (e.g., genomic scientists, veterinarians)**

##### **Task**

*Querying the database  
Submitting to the database  
Administrative functions*

##### **Users**

*Scientists, pharmaceutical industry, public  
Scientists, curators at NCICB  
NCI personnel*

#### **Animal models (species and strains) covered**

*Mouse and rat models of human cancer, mostly genetically engineered models, but database is equipped to handle xenograft models. The database has been expanded to include zebrafish and could be easily adapted to other species.*

#### **Means by which users may access the resource (e.g., downloads, browsing web site)**

*Internet <http://cancermodels.nci.nih.gov>*

*Submitters need user accounts in order to submit or edit models or comment on existing models. User accounts can be requested via the website.*

#### **Data and Content**

- Distribution and dissemination of data (if different from above)  
*Internet, <http://cancermodels.nci.nih.gov>  
API (see below)*
- Description of data schema and organization of information  
*We will provide the Enterprise Architect file upon request (file is too large to email)*
- Nature and type of data contents, especially diseases and normal and pathological conditions covered or represented  
*caMOD contains models for human cancer as well as tool strains which can be or have been used to generate cancer-bearing model.*

- Source of data contents  
*Data are submitted by scientists who generated or worked with the model in question. Data can also be extracted from literature and entered by NCICB staff with or without input from the authors.*
- Terminologies, ontologies, etc. used for indexing, annotation, curation  
*caMOD utilizes a variety of controlled vocabularies.*
  - *Murine tissue vocabulary – provide by the NCI Thesaurus, a product of the Enterprise Vocabulary System (<http://nciterms.nic.nih.gov>)*
  - *Mouse Disorders – provide by the NCI Thesaurus, a product of the Enterprise Vocabulary System (<http://nciterms.nic.nih.gov>)*  
*This disease vocabulary is a consensus product of the pathology workgroups within the Mouse Models of Human Cancers Consortium.*
  - *Many fields on the submission pages are populated by dropdown lists (controlled vocabularies). Each controlled vocabulary is accompanied by an “other” option which allows the user to submit terms missing from the vocabularies. After review by the curation team, the new terms can be added to the controlled vocabulary.*

## **Technical and Architecture**

### **Development platform and architecture**

#### User Client Software

*caMOD is a web based user interface implemented as HTML, JSP, JavaScript, and AJAX. The use of JavaScript is limited to improving usability only.*

*Java Server Pages (JSPs) are used to generate dynamic HTML based on user actions. JSPs are combinations of HTML, Java Code, and custom JSP and JSTL tags that get compiled into servlet classes.*

#### Back End

*The caMOD back end components provide access to persistence services for retrieving and storing data in the database. Hibernate, an open source tool, is used as the Object Relational Mapping (ORM) software.*

#### JBoss Application Server

*The user interface, middle tier, and backend components are run with in JBoss Java2 Enterprise Edition (J2EE) application server. JBoss is an open source application server that provides a servlet container, an EJB container, the full range of J2EE 1.4 features as well as extended enterprise services including clustering, caching and persistence.*

#### Database

*Oracle Relation Database Management System (RDBMS) software is used as the persistent store for caMOD application data. Mid-tier code accesses the database using Hibernate ORM technology and Java Database Connectivity (JDBC) API.*

### **Routine maintenance and update schedule**

*Updates as required*

### **Technical pre-requisites for use**

*Standard web browsers like Internet Explorer 6.x and Firefox 1.x and above*

### **Interoperability with other bio-informatic tools and software**

*caMOD has been designed, architected and constructed to facilitate interoperability with other systems, following caBIG™ guidelines.*

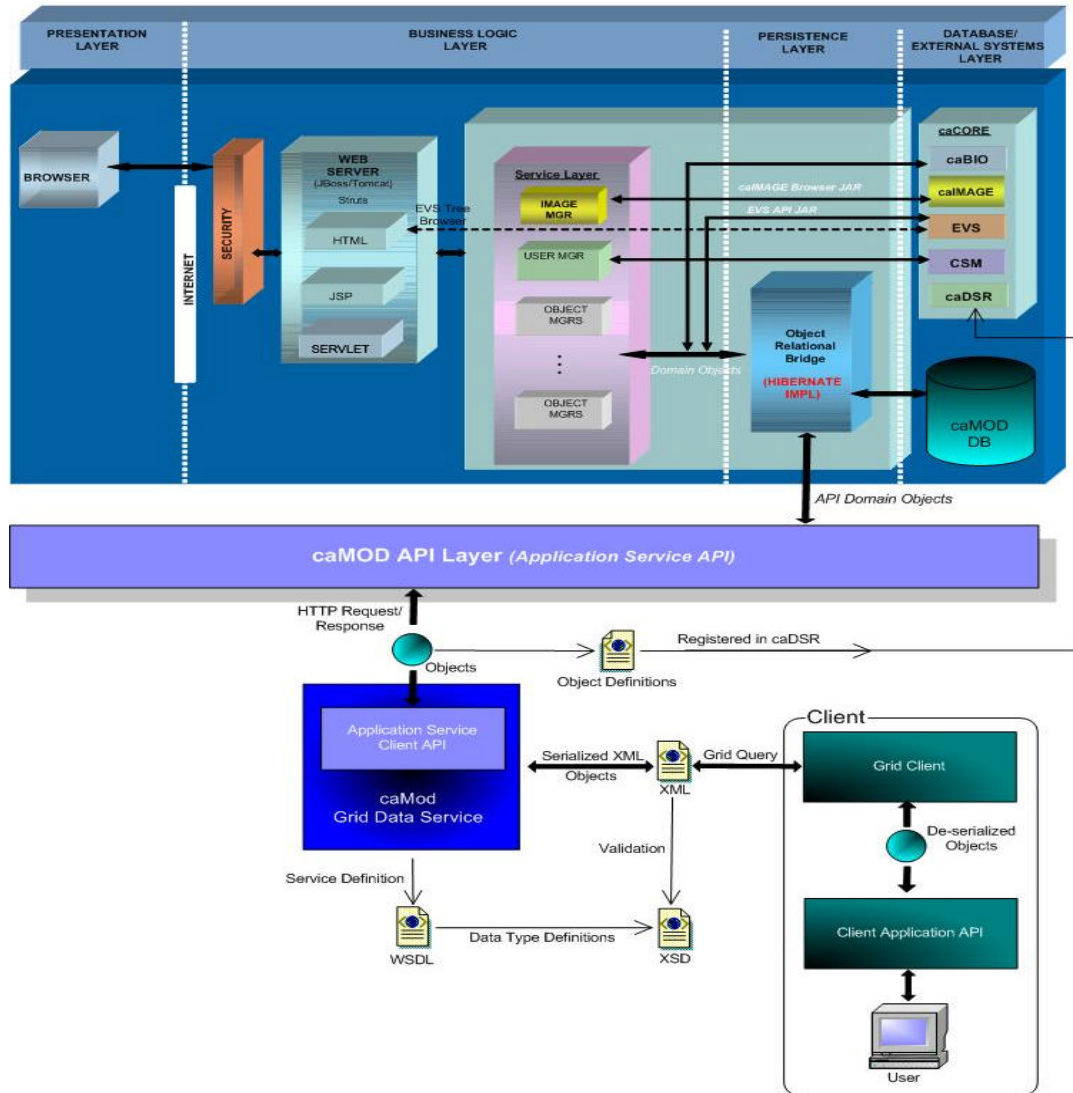
*Information Providers to caMOD:*

- *caBIO to retrieve gene info and clinical trials info through remote API*

- EVS to provide concept codes and preferred descriptions for concepts through caBIO EVS API
- PubMed
- Jackson Laboratory Resources
- NCI's Developmental Therapeutics Program
- caArray to store microarray data
- caIMAGE server to store images
- ▶ Information Consumers: caMOD provides information to other systems
  - CMAP
  - BioGopher
  - Websites such as eMice references specific models in caMOD
  - caELMIR (future)

Availability of application programming interfaces (APIs) for resource users, the functions that are provided via the APIs, the data returned and any major limitations on API use (e.g., additional licensing, dependencies, specific technical requirements)

### caMOD API Architecture



*The caMOD application has been registered as a grid node in the caGrid infrastructure. The caMOD Grid Data Service node is generated using the caGrid Toolkit. The Data Type Definitions extracted from the cancer Data Standards Repository (caDSR) registered caMOD model are used to build an XML schema or XSD that is used to define the input and output data types of the Grid services. The XSD is registered in the Global Model Exchange (GME) and used to describe the contract of the grid service and to validate the XML serialization of the API Domain Objects during data requests. Applications communicate with the caMOD Grid Data Service node following the caGrid services protocol. The caMOD API domain objects are serialized/deserialized by the caMOD Grid Data Service to and from XML adhering to the XML Schema (XSD) that is registered in the Global Model Exchange (GME)*