

Index of NCI Studies (INS) Release Notes

Newest release listed first

INS Datasets Release (v3.0.0)

NOV 26, 2024

Summary

The “INS Datasets Release” of the Index of NCI Studies (INS) is available now at [studycatalog.cancer.gov!](https://studycatalog.cancer.gov/)

This major update adds new data and features focusing on datasets generated through NCI-supported research. Initially, this includes references to NCI-supported studies registered with the [NCBI dbGaP](https://www.ncbi.nlm.nih.gov/dbgap/), though the INS will expand to other dataset resources in the future.

The NCI-supported Programs, Projects, Grants, and Publications have also been improved and expanded since the last release, with the INS now holding more than double the number of records for each.

KEY FEATURES

- Added 871 NCI-supported studies registered with the NCBI Database of Genotypes and Phenotypes (dbGaP) to the INS as the first datasets
 - Various dbGaP metadata sources were incorporated into the automated INS data gathering pipeline to pull dataset information
 - The automated dataset results were further curated for accuracy and completeness
- Added a new Explore Datasets page to allow users to search, browse, and preview dataset information
 - Card-view presentations of each dataset include basic descriptive information along with links to the INS Dataset Details page and directly to the external source (i.e. dbGaP study page)
 - Facet filters allow selection of specific Primary Diseases of interest. Results will include only datasets focused on the filter selections
 - Search bar allows for free text search that will access all text-based dataset fields and return only matching results
 - Search highlights will identify the matching term(s) in each resulting dataset
 - Export button will download a CSV with all the metadata for each dataset in the current results
- Added a new Dataset Details page to allow users to view all detailed information captured by the INS for each specific dataset
 - Includes all information displayed on the Explore Page card view in addition to more detailed information, such as Principal Investigator(s), Funding Source(s), Limitations for Reuse, and related keyword terms
 - External links to the [NCBI PubMed](https://pubmed.ncbi.nlm.nih.gov/) are provided for any relevant publications cited by the dataset

- Improved the Program Details page to present a descriptive display and external link to the program website when no research outputs are currently captured by the INS for that program
- Updated the Home Page links and graphics to integrate the datasets pages and information
- Updated the About Page to include dataset information to both the text narrative and to the data flow diagram
- Updated preview text within the header Documentation Search bar for clarity
- Added new automation practices for faster, more scalable QA testing of data
- Upgraded infrastructure and deployment pipelines to improve stability and security
- Updated software packages and resolved various security issues

Data Details

This release includes information for:

- 163 Programs
- 871 Datasets
- 1,899 Projects
- 6,376 Grants
- 35,339 Publications

Program-derived data in this release of INS were gathered on September 20, 2024, using version [2.0.0](#) of the data gathering pipeline. Dataset data were gathered and curated in September-October 2024.

[End of Release Notes for the INS v3.0.0. Release Notes for earlier updates listed below.]

INS Tooltips Release (v2.2.0)

JUNE 28, 2024

Summary

The “INS Tooltips Release” of the Index of NCI Studies (INS) is available now at studycatalog.cancer.gov! This lightweight update brings interactive tooltips to help users better understand terms used throughout the site. This release also carries several technical improvements that will aid future development.

KEY FEATURES

- Added interactive definition tooltips to most column headers across all tabs of the Explore Page
- Expanded the “Disclaimer Banner” to all pages of the site rather than only the Home Page
- Changed the fixed header to a floating header, giving users more display space and improving compatibility with other header features
- Implemented Adobe Analytics for all pages
- Improved detail of software version display in site footer
- Minor bugfixes for sorting and table display

The underlying data remains unchanged in this release.

[End of Release Notes for the INS v2.2.0. Release Notes for earlier updates listed below.]

INS Detail Pages Release (v2.1.0)

MAY 9, 2024

Summary

The “INS Detail Pages Release” of the Index of NCI Studies (INS) is available now at studycatalog.cancer.gov! This release builds upon the recent “Program Redesign” and reintroduces two page views to the site: the **Program Details Page** and **Project Details Page**. Each presents a focused view of additional details including a filtered view of the Explore Page tables relevant to the Program or Project. Every program and project within INS can now be better explored using these pages.

Example Pages

Program Details: [Fusion OncoProteins in Childhood Cancers \(FusOnc2\)](#)

Project Details: [Defining and targeting the molecular vulnerabilities of the PAX3-FOXO1 protein in rhabdomyosarcoma \(U54CA231630\)](#)

KEY FEATURES

- Implemented Program Details and Project Details Pages
- Added links to Program Details and Project Details Pages throughout the Explore page tables
- Added a new Program Cancer Type checkbox filter to the Explore Page
- Restructured data model and data gathering to separate Program Cancer Type(s) from Focus Area(s)
- Updated all INS data (programs, projects, grants, publications)
- Improved user navigation and sorting in tables
- Improved formatting of special characters within data and CSV downloads
- Improved data validation protocol for more efficient QA testing
- Upgraded infrastructure and deployment pipelines to improve stability and security

Data Details

This release includes information for:

- 83 Programs
- 757 Projects
- 3,065 Grants
- 22,199 Publications

Data in this release of INS were gathered on April 26, 2024, using [v1.1.0](#) of the data gathering pipeline.

INS Program Redesign (v2.0.0)

APRIL 4, 2024

Summary

The “INS Program Redesign” of the Index of NCI Studies (INS) is available now at studycatalog.cancer.gov! This release represents a major overhaul of the data, design, and function of the site in response to feedback gathered from initial releases.

The INS is now expanded to include 83 programs curated by the NCI’s Office of Data Sharing. These programs act as the input into a redesigned, automated data gathering process that utilizes public resources from [NIH RePORTER](#), [NIH iCite](#), and [NCBI PubMed](#) to gather projects, grants, and publications associated with each program. Features of the site have been redesigned to center on the new data design.

KEY FEATURES

Overview

- Expanded scope of NCI program curation from 2 initial programs to 83
- Redesigned automated data gathering process to find associated projects, grants, and publications for each program using stable, publicly available APIs
- Redesigned data model to accommodate new fields and connections within gathered data
- Redesigned Explore Page with all new tables, plots, and search/filter capabilities
- Removed obsolete Programs page and navigation button. Programs are now integrated within the Explore Page
- Removed Program Detail and Project Detail Pages. These will be considered for future releases
- Removed datasets, clinical trials, and patents data types. These will be considered for future releases
- Updated Home Page text and stat bar graphic
- Updated About Page documentation and design
- Updated Glossary to address new terms and changes to existing terms
- Upgraded frontend framework to support additional visualizations (Bento 4.0)
- Upgraded infrastructure and deployment pipelines to improve stability and security

Data Details

This release includes information for:

- 83 Programs

- 757 Projects
- 3,047 Grants
- 21,898 Publications

Data in this release of INS were gathered on March 13, 2024, using [v1.0.0](#) of the data gathering pipeline.

[End of Release Notes for the INS v2.0.0. Release Notes for earlier updates listed below.]

INS Minor Security Release (v1.3.0)

AUG 31, 2023

Summary

The Minor Security Release of the Index of NCI Studies (INS) is available now at studycatalog.cancer.gov. The primary focus of this release is to perform maintenance and security updates. Though important, these types of updates are generally not visible to a visitor to the site. However, we have also included updated graphics and animations on the Home Page along with various bugfixes and visual polish throughout the site.

The underlying data remains unchanged in this release.

Home page

- Added mouseover rotation animation to the INS hero image
- Improved responsive design to allow the hero image to better fit a wider range of screen sizes

General updates

- Updated the INS backend to the Bento Framework v4.10
- Updated backend packages to improve security
- Employed Search Engine Optimization strategy to help users find the site
- Improved formatting of CSV downloads and download buttons
- Improved browser tab behavior for external links in global footer

[End of Release Notes for the INS v1.3.0. Release Notes for earlier updates listed below.]

INS Core Project Remodel (v1.2.0)

MAY 4, 2023

Summary

The “Core Project Remodel” of the Index of NCI Studies (INS) is available now at studycatalog.cancer.gov. The primary focus of this release is to reorganize the data and display of the INS to center around (core) projects rather than grants, which addresses one of the major improvements identified in the initial release.

As before, we are only including outputs from projects funded at least in part by **extramural grants** associated with the **Moonshot** and **Childhood Cancer Data Initiative (CCDI)**. The task of identifying and defining programs remains a challenge to be addressed in future releases of the INS.

KEY FEATURES

Overview

- Redesigned data architecture to organize grants (formerly known as projects within the INS) into groups of grants. Those groups are now called projects and become the focal point of the INS
 - Redesigned the Project Details Page to focus on one project with all associated grants and outputs displayed. Project descriptions are pulled from the most recently awarded grant, except for the Award Amount Total fields, which display summed values from all associated grants.
 - Added Project Details Page links to all tabs of the Explore Page to quickly find all information for a project from any grant or output of that project
 - Updated Explore Page filters, Programs Page, Program Details Page, and Global Search to reflect the new roles of projects and grants
 - Updated documentation pages (About, Glossary, Release Notes)
- Added animation and updated the Stats Bar Graph image on the Home Page
- Improved data gathering date filters to reduce false positive associations
- Added 3 new manually curated CCDI P30 supplement datasets and projects

Data Details

This release includes the following data totals:

- 2 Programs
- 350 Projects
- 1,211 Grants
- 4,748 Publications

- 1,388 Datasets
- 349 Clinical Trials
- 117 Patents

New in this release of INS are several datasets shared due to CCDI P30 grant supplements to the following organizations:

- Memorial Sloan-Kettering Cancer Center (MSKCC): [phs003111.v1.p1 | 3P30CA008748-54S3](#)
 - Note: this is an additional dataset shared due to a CCDI P30 grant supplement already associated with another dataset in the INS (phs002620.v1.p1)
- University of Kansas Medical Center: [phs002529.v1.p1 | 3P30CA168524-09S3](#)
- University of Southern California: [phs002518.v1.p1 | 3P30CA014089-45S6](#)

Note that due to the indirect nature of grant supplements, these 3 projects and their outputs were gathered and added to the INS through manual curation rather than the standard automated process.

This release includes a new batch of data gathered using the automated processes. It includes improvements in the preprocessing pipeline to filter out any outputs published before the earliest associated grant was awarded. This helps to limit false positive associations.

Projects, grants, publications, datasets, and clinical trials included in this release were gathered on April 4, 2023. Due to instability in the external source from which we gather Patents (USPTO API) in April, patents included in this release were gathered on March 2, 2023.

[End of Release Notes for the INS v1.2.0. Release Notes for earlier updates listed below.]

INS Minor Release (v1.1.0)

JAN 5, 2023

Summary

The latest “Minor Release” of the Index of NCI Studies (INS) is available now at studycatalog.cancer.gov. The primary focus of this release is to introduce several key features and fixes identified in the initial release.

As before, we are only including outputs from **extramural grants** funded through the **Moonshot** and **Childhood Cancer Data Initiative (CCDI)**. We intend to address the CCDI data complications mentioned in the site disclaimer in future releases.

KEY FEATURES

Overview

- Added 4 manually curated CCDI P30 supplement datasets
- Added buttons for users to download tables as CSV files for easier ad hoc analyses
- Added INS Glossary and INS Release Notes PDFs to the main navigation menu bar
- Added “Publication Date” column to Publications tables
- Removed “NIH Percentile” column from Publications tables
- Redesigned architecture of Project Detail and Program Detail pages to improve performance and table navigation
- Many small UI improvements and bugfixes

Data Details

This release includes the following data totals:

- 2 Programs
- 1,041 Projects
- 4,098 Publications
- 1,119 Datasets
- 207 Clinical Trials
- 117 Patents

New in this release of INS are several datasets shared due to CCDI P30 grant supplement projects to the following organizations:

- Memorial Sloan-Kettering Cancer Center (MSKCC): [3P30CA008748-54S3](#)
- Oregon Health & Science University (OHSU): [3P30CA069533-22S5](#)
- Baylor College of Medicine: [3P30CA125123-14S6](#)
- University of California, San Francisco (UCSF): [3P30CA082103-21S9](#)

Note that due to the indirect nature of grant supplements, these 4 projects and their outputs were gathered and added to the INS through manual curation rather than the standard automated process.

The data in this release uses the same automated data gathering output as previous releases of INS (09/26/2022 for Projects, Publications, and Patents; 05/26/2022 for Datasets and Clinical Trials) but improves the curation to remove duplicates and other minor display issues. As mentioned above, it also includes 4 manually curated CCDI projects and their datasets.

[End of Release Notes for the INS v1.1.0. Release Notes for earlier updates listed below.]

INS Disclaimer Release (v1.0.1)

NOV 18, 2022

Summary

The latest “Disclaimer Release” of the Index of NCI Studies (INS) is available now at studycatalog.cancer.gov. The primary focus of this small release is to add additional information about challenges faced and known limitations of the initial phase of the INS.

Building connections between NCI programs, grants, and outputs is a complex task usually addressed through manual curation by portfolio analysis experts. As we work to automate the process and share this resource with the public, we would like to inform you of the known limitations within the data gathering process and information displayed by the INS. We are working on improving data gathering process and will periodically release updates which allows us to iteratively improve the INS site.

Note that the underlying data remains unchanged in this release.

Home page

- Added disclaimer banner to the header
- Changed title text and subtext to highlight known limitations

About page

- Added “Gathering Research Outputs: A Continuous Challenge” section to describe known limitations in detail

General improvements

- Improved neo4j queries on Project Detail pages to avoid potential memory overload errors

[End of Release Notes for the INS v1.0.1. Release Notes for earlier updates listed below.]

INS Initial Release (v1.0.0)

OCT 11, 2022

Summary

The National Cancer Institute (NCI) is pleased to announce the release of the **Index of NCI Studies (INS)** at studycatalog.cancer.gov.

The INS is a new website that gathers and displays publicly available information about research outputs from NCI-funded programs. The goal of the INS is to encourage an environment of FAIR (Findable, Accessible, Interoperable, Reusable) data sharing and to reinforce the NCI's commitment to transparency and accountability as the nation's leader in cancer research.

The intent of the **initial phase** of the INS is to solicit feedback for further expansion of this tool, which is limited in both scope and scale by design. The current site focuses on gathering data outputs associated with **extramural grants** from two key NCI programs: the **Cancer Moonshot** and the **Childhood Cancer Data Initiative (CCDI)**.

KEY FEATURES

Overview

This release includes the following pages and features:

- Home page
- Explore page
 - 1,037 Projects
 - 4,121 Publications
 - 1,121 Shared Datasets
 - 207 Clinical Trials
 - 117 Patents
- Programs page
 - 2 Programs
- About page
- Project Detail page
- Program Detail page
- Global Search Bar & Results page

Home page

<https://studycatalog.cancer.gov/#/home>

The Home page is the intended landing page for users viewing the site. It provides introductory information about the INS and links to other pages. Users can navigate to the Home page with a link in the main navigation bar.

Explore page

<https://studycatalog.cancer.gov/#/explore>

The Explore page provides comprehensive lists of projects and research outputs in the site and is the primary page for browsing projects and research outputs. Users can navigate to the Explore page with a link in the main navigation bar. The tables contain 5 tabs with tables grouped by project or research output type:

- Projects
- Publications
- Datasets
- Clinical Trials
- Patents

Research outputs generally have an external link to allow users to navigate to the source output directly. Projects show an internal link that take a user to the Projects view page.

The lists of projects and research outputs can be dynamically filtered by several project attributes listed below. Donut charts provide basic project metrics and are updated as filters are applied.

- PROGRAM
- DIVISION/OFFICE/CENTER (DOC)
- FISCAL YEAR
- AWARD AMOUNT

Programs page

<https://studycatalog.cancer.gov/#/programs>

The Programs page lists the programs currently included within the INS, along with links and basic metrics associated with each. Each row contains both an external link to a program's website and an internal link to the INS Program Detail page for that program. Currently, this page shows only the two programs included in

the initial release of the INS: Moonshot and CCDI. Users can navigate to the Programs page with a link in the main navigation bar.

About page

<https://studycatalog.cancer.gov/#/about>

The About page gives an overview of the INS project, data gathering process, and site functionality. Users can navigate to the About page with a link in the main navigation bar.

Project Detail page

<https://studycatalog.cancer.gov/#/project/1U01CA217885-01> (example)

The Project Detail page gives detailed information concerning a specific project, categorized into sections for Basic Info, PI and Organization, and Award information. In addition, the Project Detail page contains tables listing research outputs associated with the parent project. Tables for different types of research outputs can be accessed by clicking on the corresponding tab (Publications, GEOS, SRAs, dbGaPs, Clinical Trials, or Patents). The tables accessible within this page are a subset of the data accessible via the Explore page.

Program Detail page

<https://studycatalog.cancer.gov/#/program/Moonshot> (example)

The Program Detail page gives detailed information concerning a specific program. This includes sections for basic descriptive information, external links, and donut charts of selected metrics. In addition, the Program Detail page contains a table of all associated projects. The table accessible within this page is a subset of the data accessible via the Explore page.

Global Search Bar & Results page

<https://studycatalog.cancer.gov/#/search/>

The Global Search Bar accepts searches of Project details and About page contents and returns the relevant Project Detail page or About page section via the Results page. The search function currently only accepts exact prefix matches. For example, searching for project ID “1U01CA217885-01” will successfully return the project, but “U01CA217885-01” will not. Improving this search functionality will be a priority for future releases.