Epic Cosmos | Overview of Cosmos

With participation from hundreds of organizations combining EHR data from hundreds of millions of patients, we can improve the health and lives of people everywhere. Cosmos makes that possible. It brings together data from across the Epic community to form the universe's largest database of EHR patient information, which you can use to:

- **Gain insights at the point of care.** Best Care Choices for My Patient will help your clinicians make key clinical decisions—such as which medication to prescribe or whether to perform surgery—by looking at research insights powered by Cosmos and showing what's worked best for similar patients.
- **Accelerate the pace of healthcare innovation.** You can draw upon the collective data of the Epic community to analyze health information and create knowledge at an unprecedented scale.
- **Collaborate with peers.** With the Look-Alikes feature, Cosmos connects clinicians caring for similar patients with rare characteristics—so they can collaborate and learn from each other.

For the latest information about the Cosmos database, including patient and encounter counts, contributing organizations, and published studies, refer to Cosmos.Epic.com.

How Does Cosmos Work?

Organizations choose to participate in Cosmos and contribute a HIPAA-defined limited data set for all patients to the centralized Cosmos database. Every Cosmos user gets access to the data set that continues to grow over time. Users can explore Cosmos with SlicerDicer, Epic's data exploration tool, or with sophisticated data science tools available within a user-specific virtual machine.

We collect feedback from the Epic community about use cases for Cosmos. Based on that feedback, the definition of a limited data set, and guidance from the Cosmos Governing Council, we define the Cosmos data set.

Cosmos Data Set

The following table lists the core Cosmos data domains and the key elements included within those domains.

Domain	Key Elements
Allergies	Allergen, reaction, reaction type, severity

Birth History	Delivery method, gestational age, birth count, birth order, birth weight, discharge weight, length, head circumference, APGAR, length of stay, type of nourishment
Cancer Staging	Stage, stage group, type or site, classification, method
Care Teams	Provider specialty
Demographics (Limited Data Set)	Age (age in days for patients under 9 weeks; age in weeks for patients between 9 weeks and 6 months; and age in months for all others), legal sex, gender identity, sex assigned at birth, race, ethnicity, country, state, county, 5-digit ZIP, patient status, date of death, cause of death, spoken and written language, marital status
Diagnoses	Diagnosis, problem list diagnosis, billing diagnosis, ED diagnosis
Encounters and Admissions	Encounter type (outpatient, admission, ED, HOV), department specialty, means of arrival, discharge disposition, ED disposition, reason for visit/chief complaint, insurance/financial class, ED acuity, ICU stays, ED event timing data, admission and discharge timing data
Family History	Relationship (including child-birthing parent linkages), relation legal sex
Geographic Factors	RUCA, Social Vulnerability Index (SVI), Area Deprivation Index (ADI)
Immunizations	Immunization, route
Lab Component Results	Component, result value, abnormal flag, LOINC
Medications	Medication (prescription, hospital or clinic administration, patient-reported)
Patient Experience Data	MyChart usage, patient-reported outcomes (questionnaires, questions, responses)
Procedures	Clinical procedures, billed procedures, genomic variant results
Social History	Tobacco, alcohol, substance use, birth control, SDOH
Transplants (Liver and Kidney)	Organs, progression dates, ischemic times, donor serologies (HCV, EBV, and CMV), donor types
Vitals	Blood pressure, pulse, temp, respirations, SpO2, height, weight, BMI, head circumference

In-Progress Data Types

We've starting pulling data from the following data types into the Cosmos database, but it isn't available for users in the Cosmos portal quite yet.

- Breast imaging data
- Expanded OB history (pregnancy outcomes, labor management, pregnancy complications)
- Genomic variants
- Infections, including hospital-acquired infections

- Medicare fee schedules
- Microbiology results (organism, antibiotic, sensitivity)
- Mixture medications
- Patient-entered questionnaires
- Research studies
- Risk scores
- SmartData elements (record types: patients, encounters, orders, notes, history, problems, cancer staging)
    - We add reporting tools to the portal for SmartData elements based on specific use cases. If you have a use case you want supported, contact your Epic Cosmos representative.
- Ventilator use and intubation

To learn more about what data types we're working on and what we have planned for the future, refer to the Cosmos roadmap.

De-Identified Data Set

A de-identified version of the Cosmos data set is available for users who access the data set directly using data science tools in Cosmos.

We worked with Privacy Analytics Inc., a company specializing in data privacy solutions, to determine a transformation strategy in accordance with the Expert Determination Method of De-identification described in the HIPAA Privacy Rule at 45 CFR 164.514(b)(1).

The de-identified data set supports strong patient privacy protections within Cosmos while maximizing the utility or the data set for research purposes.

In the de-identified data set, we transform the data in the following ways:

- Suppression of the following data elements:
    - State
    - County
    - ZIP code
    - Marital status
    - Preferred language
- Suppression of the following social determinants of health (SDOH) responses:
    - Highest education level
    - Income/financial status
    - Food worry
    - Food scarcity
- Date-shifting the following, using a random offset in the range of 0-30 days. Each patient's date is shifted by the same offset to ensure consistency among their dates:
    - Patient dates, including date of birth and death date

    - After the date shifting, patients 18 and older have their date of birth generalized to January 1 of their birth year
- Generalization of country into two categories: United States or non-United States
- Addition of randomly chosen noise (from a Gaussian distribution) to SVI values
- Generalization of RUCA codes to the following intervals: 1-6, 7-10, or 99
- Suppression of diagnosis codes (ICD-9, ICD-10, SNOMED) relating to events that could be publicly knowable, such as natural disasters and military operations
- Suppression of information about the source organization that submitted the data

Data Transfer

The Cosmos engine is built on capabilities derived from Care Everywhere to exchange patient records using CDA documents. Cosmos CDA documents include only a limited data set, not full PHI. The Cosmos exchange works similarly to the organization-to-organization patient exchanges in Care Everywhere already in use across the Epic community.

You don't need an additional data warehouse, ETL, or i2b2 instance to participate in Cosmos. In most cases, no additional servers are required. Before you submit data, Epic will work with you to perform a Cosmos system impact assessment to understand, plan for, and minimize impact to your overall system performance.

Longitudinal Records

After you sign up for Cosmos and complete the initial setup, the Cosmos engine submits historic patient data in configurable increments to Cosmos.

Updates to Records

When patient or encounter data changes, your system flags and transmits the updates to Cosmos.

Cross-Organization Patient Matching

Cosmos enables cross-organization patient matching without the need for patient demographics through patient links established in Care Everywhere. These links (Care Everywhere IDs) are masked using a SHA-256 hashing mechanism when sent to Cosmos. This obscures the IDs in an irreversible way. As new links are established through Care Everywhere, Cosmos identifies and merges cross-organization duplicates.

Excluding Protected Data

Cosmos is built on a flexible infrastructure that can accommodate different state-specific regulatory and consent requirements for data. During the Cosmos implementation process, your Epic representative will work with you to configure your Cosmos data submission as necessary to comply with state and local laws.

Mapping

Cosmos uses industry-standard terminologies when available. As with any data sharing program, it's necessary to map data to these terminologies so that every organization's data is

standardized and uniform when aggregated. We provide mapping tools and processes to help you with additional Cosmos-specific mapping. Cosmos reuses mappings already completed for projects such as regulatory reporting or Happy Together. Conversely, mapping that you do for Cosmos will reduce the mapping needed for other interoperability projects.

Mapping in the Foundation System

Starting with the May 2020 version of the Foundation System, the majority of data domains are pre-mapped to the required standards for Cosmos. This means that if you're a newly implementing organization that takes a cut of the Foundation System, much of your Cosmos mapping is already complete out of the box. You just need to complete additional mapping for lab result components and some procedures, as well as mapping for any custom values not included in the Foundation System that you decide to add—these processes are covered by your implementation so you can be ready to participate in Cosmos once your system is live.

The Cosmos Portal

Initially, users explore Cosmos using SlicerDicer over a secure web portal. The SlicerDicer instance used for Cosmos is hosted by Epic, so you don't need to be live with SlicerDicer or Caboodle (the underlying data warehouse) at your own organization.

SlicerDicer enables complex query and visualization capabilities, while creating a layer of separation between users and the raw line-level patient data. Query results are returned through different visualizations, counts, and percentages. SlicerDicer provides broad and general access to users without the need for individual study-level approval each time a Cosmos user wants to explore the data.

The Cosmos portal also includes data science tools, such as Python and R, that certified users can access through a user-specific virtual machine. The data science tools connect to the de-identified Cosmos data set described above.

Additional functionality, such as the Look-Alikes and Best Care Choices for My Patient features described above, will become available to participants as well.

Data Security

Cosmos is built in a secure way to protect patient privacy and data security by enforcing strict control on how data is collected, transmitted, stored, and accessed.

The points below highlight security measures at high level. If you're already signed up for Cosmos, you can review the Cosmos Security and Privacy document for more details.

Transmission and Storage

- No direct patient identifiers are stored in Cosmos.
- Data is transmitted using the secure Care Everywhere network.
- Cosmos is hosted by Epic in our secure data center.
- Data cannot be exported outside the Cosmos environment.

Access Control

- Participating organizations review their Cosmos users on a recurring basis. This provides regular opportunities to renew access for active users and end access for those who should no longer have it.
- All Cosmos users must have a current association with a Cosmos-participating organization.
- Users are provisioned at the request of participating organizations.
- Epic logs all logins and all queries executed within Cosmos for audit purposes. Query logging is performed using logging built into SlicerDicer as well as native logging available through SQL Server.

Data Set Protection

- Users electronically sign a HIPAA-required data use agreement as they first gain access to Cosmos.
- Data is shown in SlicerDicer at a summary level only, meaning users cannot drill down into line-level data (the data elements specific to one patient or encounter). This ensures the output is de-identified.
- SlicerDicer masks low cell counts when queries return fewer than 11 results. For example, the output appears as "11 or fewer" rather than the actual number. This threshold was selected to match CMS's recommendation on the Limited Data Set Data Use Agreement form (section 8a) to not show cell counts lower than 11. It was validated by Cosmos Governing Council as well.
- Users cannot export data.

How Does Cosmos Participation Work?

When you sign up for Cosmos, you sign a license amendment, which includes the Cosmos Rules of the Road as an attachment. The Rules of the Road defines the expectations of all participants. Changes to the Rules of the Road can be proposed by organizations through the Cosmos Governing Council, voted on by all Cosmos participants, and approved by Epic.

Governance

The Cosmos Governing Council is a community-elected 13-member group of representatives from Cosmos Participant organizations (expanding to 15 members in spring 2023). Members of the Council provide strategic guidance on Cosmos functionality and help define expectations of use for the community. Members serve voluntary 36-month staggered terms, with elections each spring.

Defining Cosmos Users

When your organization participates, you get access to Cosmos and get to define which users associated with your organization get access (in conformance with the Rules of the

Road). Typical Cosmos users are physicians, non-physician practitioners, and researchers, but there is no specific role requirement. Cosmos users must be either of following:

- Employees of organizations participating in Cosmos or employees of Affiliates of participating organizations (see the definition of "Affiliate" in your license agreement).
- Clinicians (RNs, MDs, PAs, PhDs from a medical field, med students) from an organization participating in Cosmos or an Affiliate of a participating organization.

In either case, the user must have an Epic login associated with the organization participating in Cosmos. Users electronically sign a HIPAA-required data use agreement as they first gain access to Cosmos.

To ensure that only current users have access to Cosmos, participating organizations review their Cosmos users on a regular basis to renew access for active users and end access for users who no longer should have it.

Third-Party Access

Third parties, such as consultants, vendors, and outsourcing firms, are not eligible for access to Cosmos unless recommended by the Cosmos Governing Council, approved by a majority of Cosmos participants, and agreed to by Epic.

Selling Cosmos data or access to Cosmos is prohibited by the Cosmos license amendment and Rules of the Road.

Participation Costs

Cosmos has no associated license or maintenance fees.

Using Cosmos for Research

You can use Cosmos for both non-sponsored and sponsored research use cases, as long as the research aligns with the permitted uses of Cosmos data. These are defined in the Rules of the Road as *using Cosmos to improve the understanding of the causes, treatment, and prevention of disease, to advance healthcare knowledge*.

Non-Sponsored Research

Some examples of non-sponsored research use cases include:

- **Following a hunch.** Have a hunch or an idea? You can use Cosmos to quickly investigate a hunch in SlicerDicer and return preliminary results that you don't plan to publish.
- **Publishing non-sponsored research.** Have an idea for a study where Cosmos has the data you need? Cosmos allows adding inclusion/exclusion criteria and basic statistical analysis on the data set that you could use to submit your research for publication.
- **Completing preparatory work for research.** Planning to do research to apply for a grant? You can use Cosmos to do your prep research for a future grant.
- **Responding to media inquiries.** Is your local paper asking for some information for a story? You can use Cosmos to supply data on recent trends that can help support a

story. Please remember to follow your organization's media communication policies prior to sharing.

Third-Party Sponsored Research

You can use Cosmos for third-party sponsored research. In fact, you might find that your participation in Cosmos opens up new opportunities for sponsored research.

If you use Cosmos for research that is sponsored by a third party, the Cosmos Governing Council and Epic will review your use case to validate that it meets the Cosmos mission to create generalizable medical knowledge. For any sponsored research, a study fee will be applied based on your study budget and expected use of Cosmos. Government-sponsored research will have a lower study fee because it is most likely to be consistent with the do-good mission of Cosmos by benefiting public health and social care initiatives. The funding is used to keep Cosmos running and available to the research community.

Research Project Tracking

Cosmos users create projects in the portal to represent their research work. When initially creating a project, users enter information about their plans and answer questions designed to help determine whether the project is funded. For sponsored projects, information is sent to the Cosmos Projects group to track the use case and potentially trigger a discussion with your organization's team to confirm Cosmos has the appropriate data and will support the use case.

Some projects might also be reviewed by the Cosmos Governing Council to ensure they meet the spirit of the Rules of the Road.

Publishing Cosmos Findings

When Cosmos users are ready to share findings from Cosmos, they use the Publication Checklist in the Cosmos portal, which includes a requirement to notify Epic of publications or presentations based on Cosmos Data. We also provide an attribution statement for use in publications.

Cosmos As a Promoting Interoperability Registry

Some organizations have decided to use Cosmos as one of their registry choices for the Public Health and Clinical Data Exchange objective for Promoting Interoperability reporting. We aren't aware of a process to confirm that any particular registry definitively counts for Promoting Interoperability (Meaningful Use) or MIPS. Based on the guidance we've seen, Cosmos appears to meet the expectations for the clinical data exchange registry measure. The ONC has not defined a technical electronic submission standard for the type of data submitted to Cosmos. We encourage you to read over CMS's guidance and make your own determination as to whether Cosmos meets the requirements.

Getting Started with Cosmos

Here's what you can do to prepare for Cosmos:

- **Sign the Cosmos license amendment including the Rules of the Road.** It might take time to work through your organizational approvals. Start early!
- **Identify a Cosmos point person.** We recommend identifying someone to own the success of getting Cosmos up and running at your organization. The point person should be responsible for staying in the loop on Cosmos communications from Epic, working directly with Epic support when necessary, and helping coordinate the governance, mapping, and general setup efforts at your organization.

Additionally, here's a look at the setup tasks you'll undertake to get Cosmos live. You can find a full list of detailed tasks in the Cosmos Setup and Support Guide.

- **Complete data mapping in Epic.** Data mapping is largest setup component for Cosmos. To complete mapping, you work with your Epic support team to streamline the effort with standard tools and processes. Mappings you do for Cosmos will also work for Happy Together. Epic will help you determine a mapping estimate for your organization.
- **Complete the configuration to turn Cosmos on.** Along with mapping, there is some additional system setup and validation necessary to get Cosmos up and running.
- **Give users access to Cosmos.** Your Cosmos point person should work with your organization's decision makers to decide who should get access to Cosmos. This group of users will receive access once your organization is live with Cosmos.

Contact us at cosmos@epic.com to get started.