

Product or feature is in a pre-release state and might change or have limited support. For more information, see the [product launch stages](#) (/products/#product-launch-stages).

Disclaimer: This operation processes data using a mixture of rules-based and heuristic methods. Results may differ between datasets and datasets due to factors like input data quality/consistency, heuristic-based algorithms, and others. This feature is not guaranteed to satisfy any specific legal, regulatory, or compliance requirements, including requirements for the de-identification of data. It is the user's responsibility to ensure that they set the appropriate configuration parameters for the operation and evaluate the end result to determine whether it is acceptable for their use cases and any legal, regulatory, or compliance requirements they may have.

De-identification is the process of removing identifying information from data. The Cloud Healthcare API detects sensitive data in [DICOM instances](#) (/healthcare/docs/how-tos/dicom-deidentify) and [FHIR resources](#) (/healthcare/docs/how-tos/fhir-deidentify), such as protected health information (PHI), and then uses a de-identification transformation to mask, delete, or otherwise obscure the data. De-identification has multiple uses cases: for example, you might want to de-identify data before analyzing or sharing it.

De-identification works at the following levels:

- At the dataset level. De-identification occurs on all data in DICOM stores and FHIR stores in the dataset. If a dataset contains both DICOM instances and FHIR resources, you can de-identify all of the instances and resources at the same time.

To de-identify sensitive data at the dataset level, call the Cloud Healthcare API [datasets.deidentify](#) (/healthcare/docs/reference/rest/v1beta1/projects.locations.datasets/deidentify) method.

- At the FHIR store level. De-identification occurs on all data in a specific FHIR store in a dataset.

To de-identify sensitive data at the FHIR store level, call the Cloud Healthcare API [fhirStores.deidentify](#) (/healthcare/docs/reference/rest/v1beta1/projects.locations.datasets.fhirStores/deidentify) method.

- At the DICOM store level. De-identification occurs on all data in a specific DICOM store in a dataset.

To de-identify sensitive data at the DICOM store level, call the Cloud Healthcare API `dicomStores.deidentify`

(`/healthcare/docs/reference/rest/v1beta1/projects.locations.datasets.dicomStores/deidentify`) method.

De-identification does not impact the original dataset, FHIR store, DICOM store, or the original data. Depending on how you configure the de-identification, the operation behaves as follows:

- If you are de-identifying data at the dataset level, de-identified copies of the original data are written to a new dataset called the *destination dataset*.
- If you are de-identifying data at the DICOM- or FHIR store level, de-identified copies of the original data are written to a new or existing DICOM or FHIR store in an existing dataset. The new DICOM store and FHIR store are called the *destination DICOM store* and *destination FHIR store*, respectively. The operation creates the destination data store if the data store does not already exist.

Note that the source dataset, FHIR store, or DICOM store and the destination dataset, FHIR store, or DICOM store must reside in the same Google Cloud project. De-identifying data across multiple Google Cloud projects is not supported.