i2 TextChart Standard

This is the placeholder topic that introduces i2 TextChart Standard, and the components that are available with this product.

Installing i2 TextChart

This topic describes how to install and activate i2 TextChart.

Before you begin

Important: If you are upgrading from an earlier version of i2 TextChart, follow the instructions in Upgrading i2 TextChart before you complete the procedure below.

Prerequisite software

The release notes contain a full list of the required software for i2 TextChart. You *must* install i2 Analyst's Notebook before you install i2 TextChart.

Optional configuration

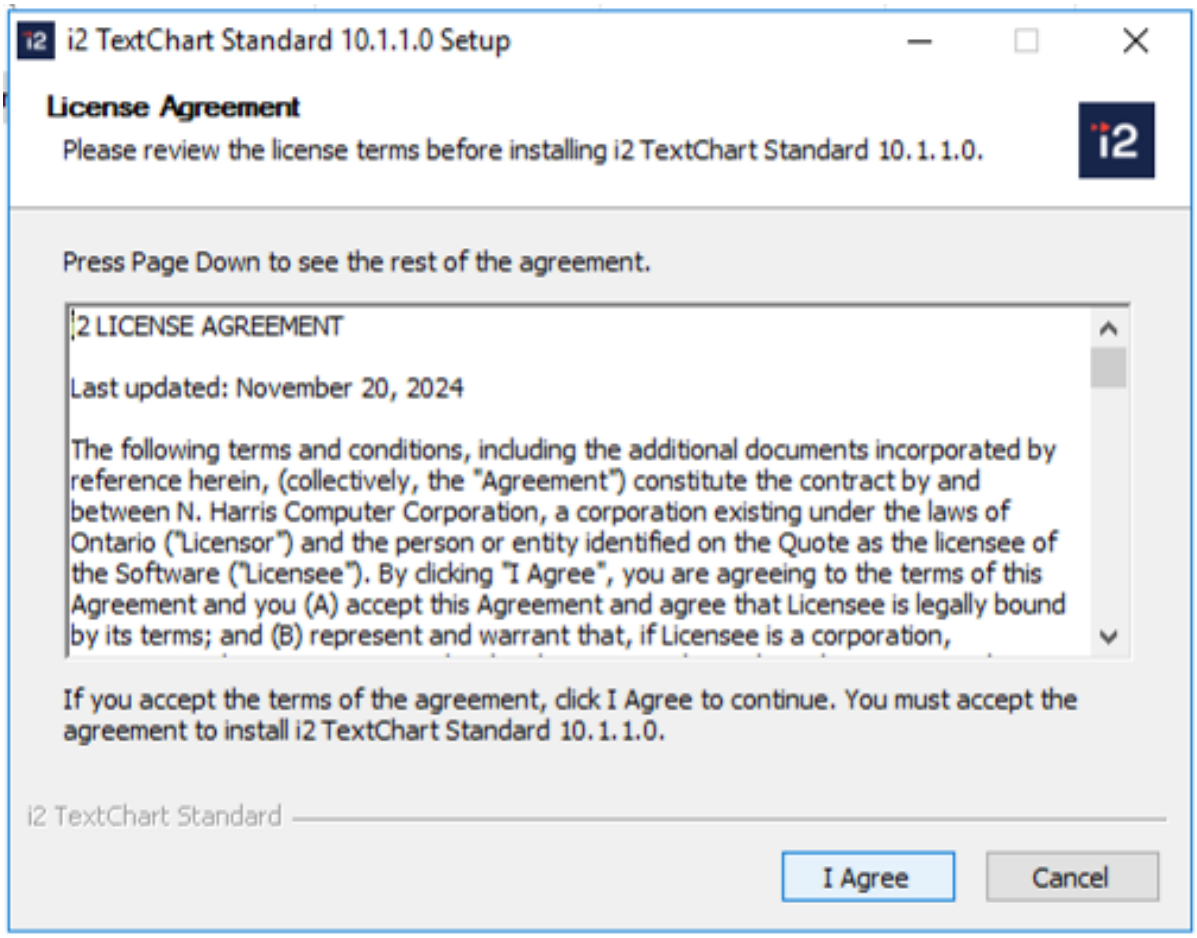
i2 TextChart is supplied with an embedded database for local workspace storage. Optionally, you can replace the embedded database with Microsoft SQL Server 2014 or later.

Procedure

The following steps describe how to install i2 TextChart.

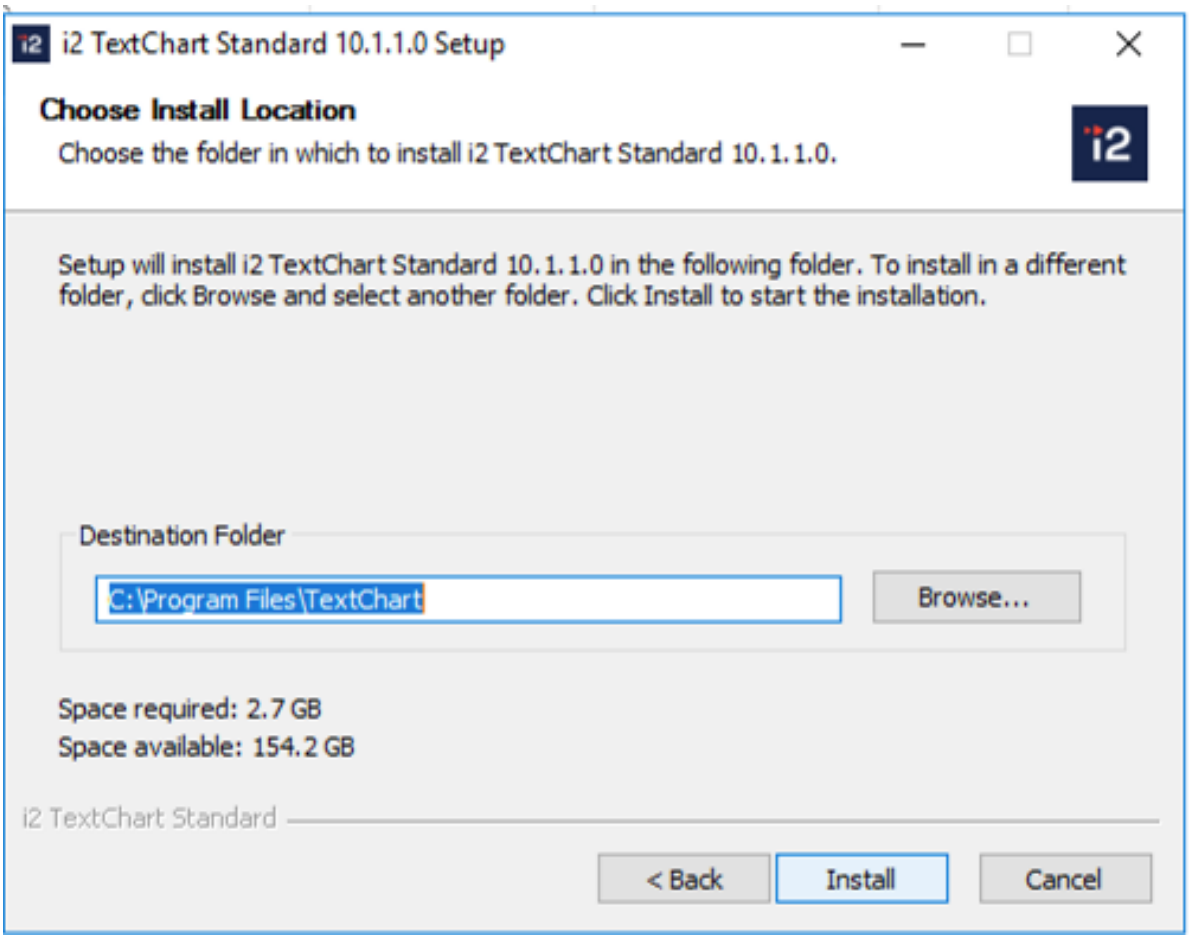
1. Extract the product files from your downloaded distribution.
2. Using Windows Explorer, browse to the root of the distribution and run the i2 TextChart Installer. Its name will contain the word "Standard" or "Premium", depending on which edition of i2 TextChart you're installing.
3. Follow the prompts. You must accept the license agreement to proceed. Click **I Agree**:

© N. Harris Computer Corporation



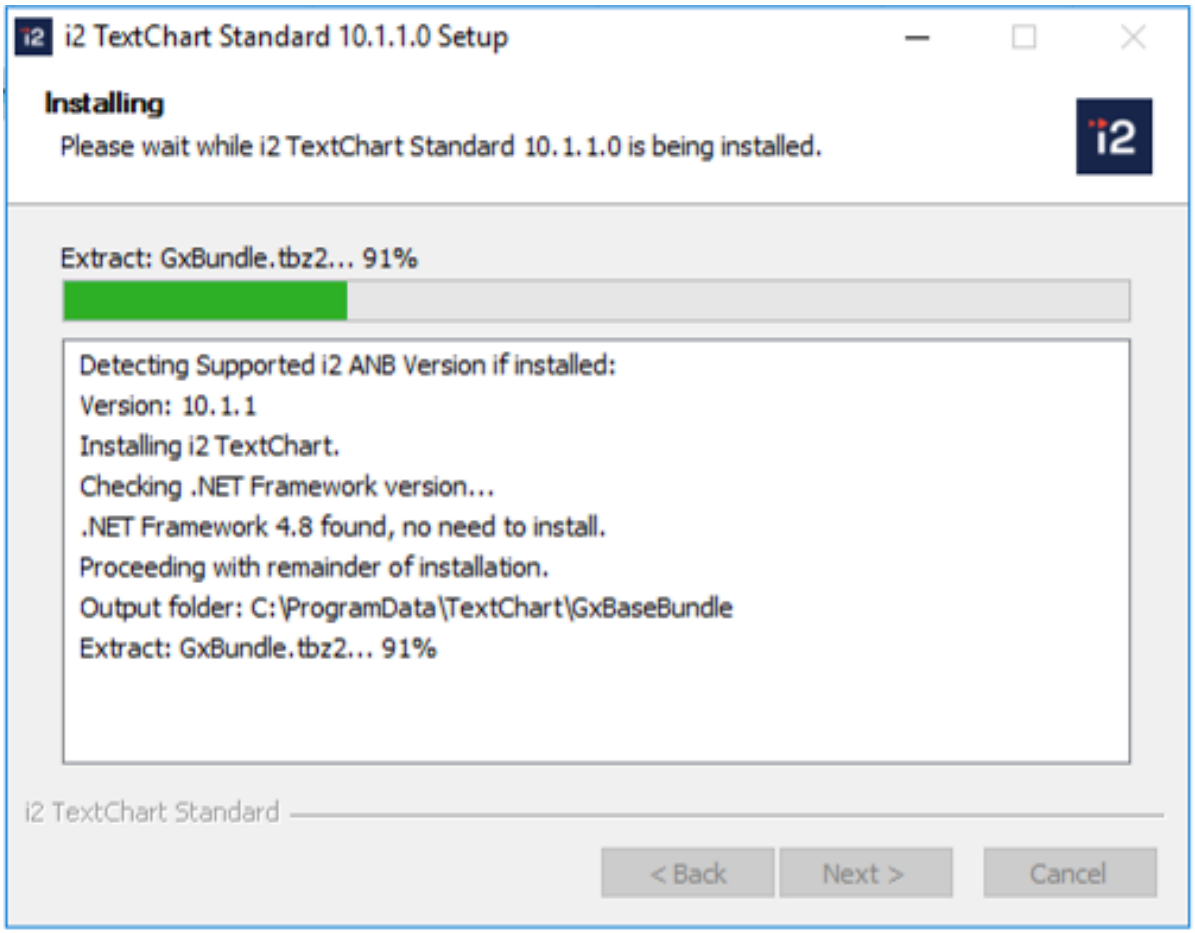
4. Choose the installation location and click **Install**:

© N. Harris Computer Corporation



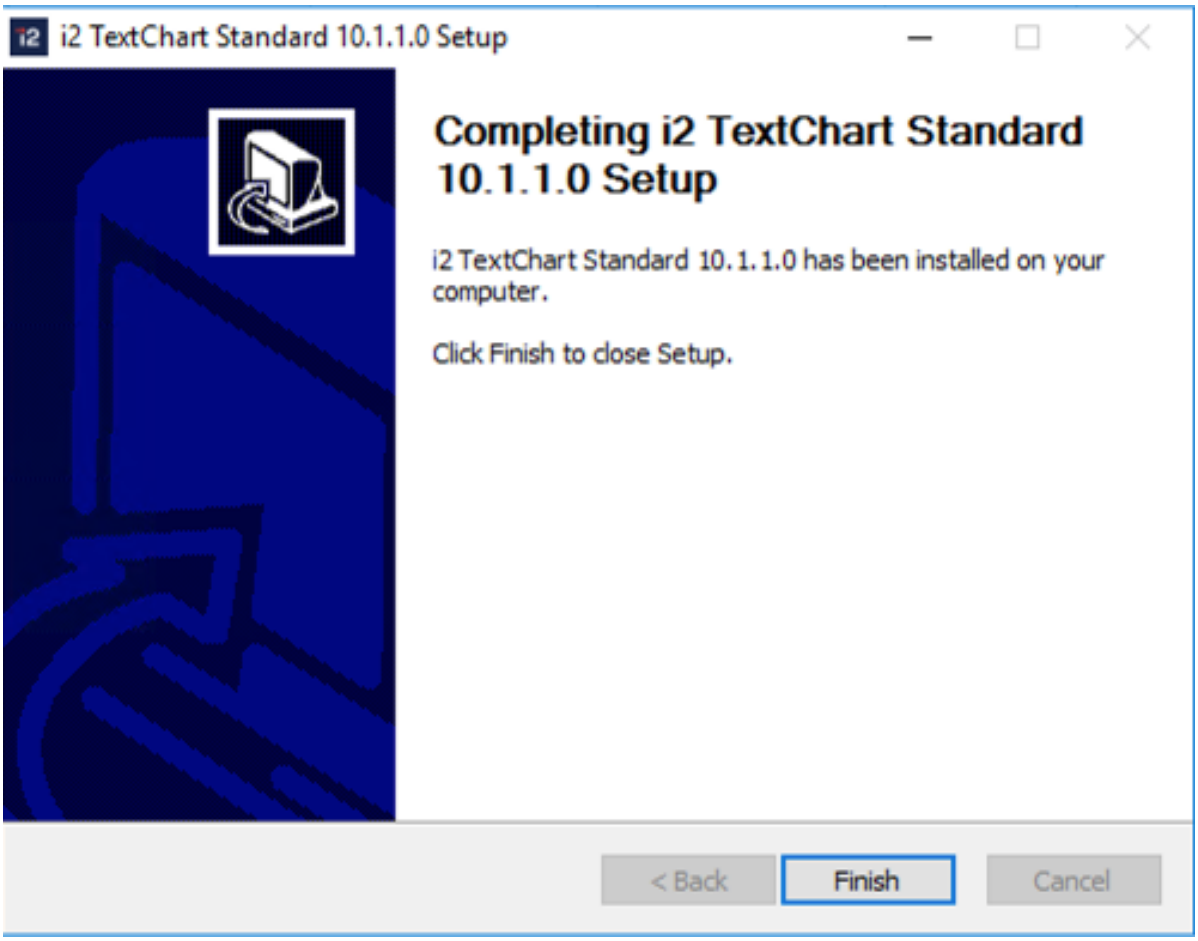
5. When the installer is complete, click **Next**.

© N. Harris Computer Corporation



6. When setup is complete, click **Finish**.

© N. Harris Computer Corporation



Installer Command-Line Parameters

The installer supports optional command-line flags that control how the application is installed. By default (no flags), the installer uses **command start/stop mode** with a **non-roaming (shared) profile**.

`/ROAMING`

Installs the application in roaming profile mode. Configuration and data are stored in the user's roaming `%APPDATA%` directory rather than the shared `C:\ProgramData\` location.

Use this when:

- Users have roaming profiles (e.g., VDI environments)
- The application data needs to follow the user across machines

`/SERVICE`

Installs the application as a Windows service (using Apache Commons Daemon). The service runs at the system level.

Use this when:

- The application should start automatically with Windows

© N. Harris Computer Corporation

- A system-level service is preferred over user-initiated start/stop

Incompatible combination

`/SERVICE` and `/ROAMING` **cannot be used together**. The Windows service runs at the system level and cannot access a user's roaming profile directory. If both flags are passed, the installer will:

- **Interactive mode**: Display an error dialog and abort.
- **Silent mode**: Write an error to `install_error.log` in the installer directory, set exit code 1, and abort.

Installing i2 TextChart Silently

The following steps describe how to install i2 TextChart silently:

1. Extract the product files from your downloaded distribution.
2. Using Windows Command Line Terminal, navigate to the root of the distribution and run the i2 TextChart Installer by issuing following command. Its name will contain the word "Standard" or "Premium", depending on which edition of i2 TextChart you're installing. Note that Analyst's Notebook is closed while installing the TextChart Text Analytics.

   ```
   "i2 TextChart Standard-X.X.X.X-Installer.exe" /S
   ```

   Note: If you need to uninstall i2 TextChart Text Analytics silently, you can do so by issuing following command. Note that Analyst's Notebook is closed while uninstalling the TextChart Text Analytics.

   ```
   Uninstall.exe /S
   ```

   Uninstall.exe should be available in installation directory, `C:\Program Files\TextChart` for example.

Installation Examples

```
:: Default install (command mode, non-roaming)
"i2 TextChart Standard-X.X.X.X-Installer.exe"

:: Roaming profile install (command mode)
"i2 TextChart Standard-X.X.X.X-Installer.exe" /ROAMING

:: Service install (non-roaming)
"i2 TextChart Standard-X.X.X.X-Installer.exe" /SERVICE

:: Silent default install
"i2 TextChart Standard-X.X.X.X-Installer.exe" /S

:: Silent roaming install
"i2 TextChart Standard-X.X.X.X-Installer.exe" /S /ROAMING

:: Silent service install
"i2 TextChart Standard-X.X.X.X-Installer.exe" /S /SERVICE
```

Defaults Summary

Setting	Default (no flag)	With flag
Profile mode	Non-roaming (`C:\ProgramData\...`)	`/ROAMING` = User roaming `%APPDATA%\...`

© N. Harris Computer Corporation

Setting	Default (no flag)	With flag
Run mode	Command start/stop	`/SERVICE` = Windows service

What to do next

When i2 TextChart is successfully installed, activated, and configured, the plug-in launches automatically when you open i2 Analyst's Notebook.

Note: If you need to stop i2 TextChart Text Analytics, you can do so through the Service Manager, which is available from the **i2 TextChart** menu in the Analyst's Notebook ribbon. To restart i2 TextChart Text Analytics, you must then restart Analyst's Notebook.

Configuring i2 TextChart

i2 TextChart is ready to use as soon as you install and activate it. However, to maintain and enhance its capabilities, TextChart also provides a range of options for configuring and updating it.

Changing the supported file extensions

The TextChart plug-in for Analyst's Notebook supports processing documents in a wide range of *formats* including plain text, PDF, and those that productivity applications generate. It uses file *extensions* to filter the documents that it presents for import, and you can change the configuration to display more or different files.

TextChart stores its list of file extensions in the file at `C:\ProgramData\TextChart\conf\filetypes.properties`. If an extension appears in this file, then documents with that extension appear in the Document View. By default, TextChart presents documents with the following extensions:

```
txt
rtf
pdf
doc
docx
pptx
```

The default `filetypes.properties` file also specifies that TextChart should present files that have no extension for import.

If your organization uses extensions that are not in the default list, you must add them to it. For example, you might have older Microsoft Powerpoint files with the `.ppt` extension, or plain text files that use a custom extension.

After you make changes to the list of supported file extensions, reinitialize the TextChart plug-in by closing and reopening the Analyst's Notebook application.

Changing the local database

By default, the standard edition of i2 TextChart uses its embedded database for storing the information that it generates during analysis. Optionally, you can configure TextChart standard edition to use a local Microsoft SQL Server database instead.

© N. Harris Computer Corporation

Before you begin

Choosing a SQL Server database in place of the embedded one enables you to inspect its contents directly. TextChart supports SQL Server 2014 and later, including the Express edition.

To configure TextChart to use a local SQL Server database, you use the Text Analytics Configuration application, which you also used to provide TextChart with a license key during installation.

To run the application after installation, you can start it from the Windows Start menu, or by launching the executable at `C:\Program Files\TextChart\RosokaConfiguration.exe`.

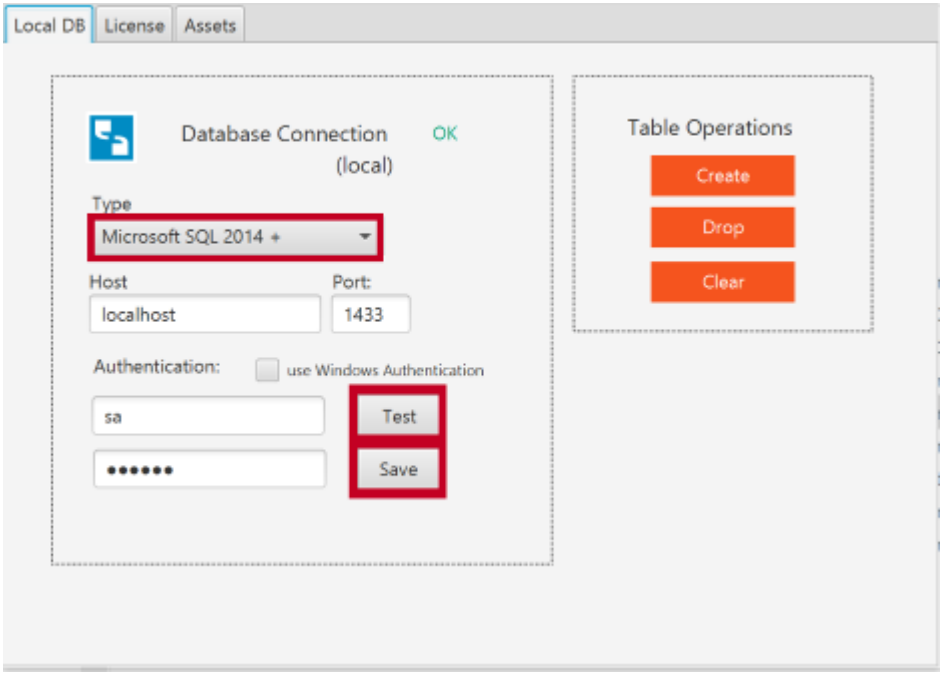
Procedure

Assuming that you have already set up a SQL Server database for TextChart to use, the procedure for connecting to it is as follows.

Note: If you want to create a database account for TextChart to use, or to change the port on which SQL Server accepts connections, see the instructions in Database configuration.

1. On the **Local DB** tab of the TextChart Text Analytics Configuration application, select your database from the **Type** drop-down, and ensure that the **Host** and **Port** details are correct.

   Note: For the standard edition of TextChart, the host must be the same workstation that TextChart and Analyst's Notebook are installed on.
2. In the **Authentication** fields, provide the user name and password of a database account with administrator privileges.



3. Click **Test** to verify that the information you provided is correct. If the test is successful, click **Save**.

© N. Harris Computer Corporation

Note: If you see the "Driver not Present" error message, you need to add the appropriate JAR file to the `C:\Program Files\TextChart\Drivers` directory, relaunch the configuration application, and try again.

Drivers for Microsoft databases are included, drivers for MySQL are available from the MySQL website, and drivers for Db2 are available as part of your Db2 installation.

What to do next

If you need to change the configuration of the local database after you've begun to use i2 TextChart, you must close Analyst's Notebook and ensure that the TextChart application is no longer running before repeating the above procedure.

Note: After you change the configuration to use a different database from the one you've been using, you might need to re-create your collections and reprocess the documents they contained.

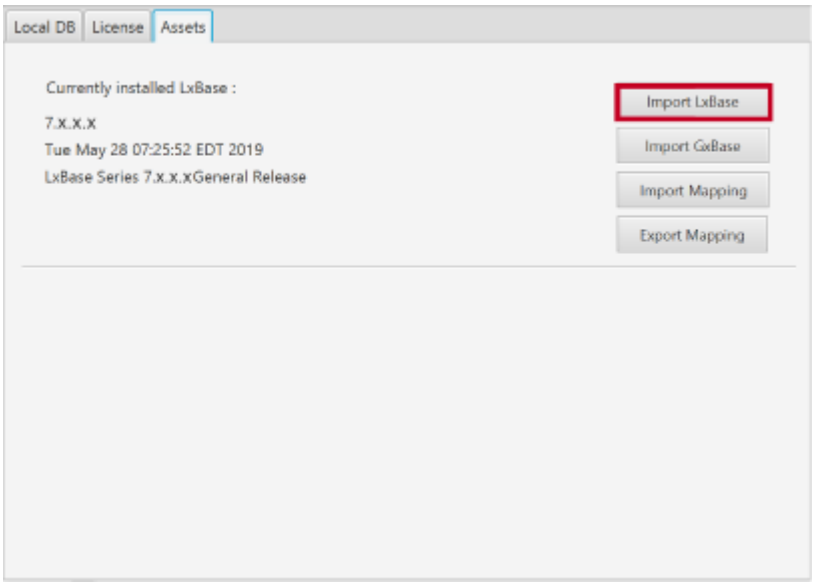
Updating an LxBase

In i2 TextChart, an LxBase is the collection of linguistic rules and dictionary entries that govern text extraction. From time to time, i2 publishes updates that you can apply to the LxBase in your installation.

Procedure

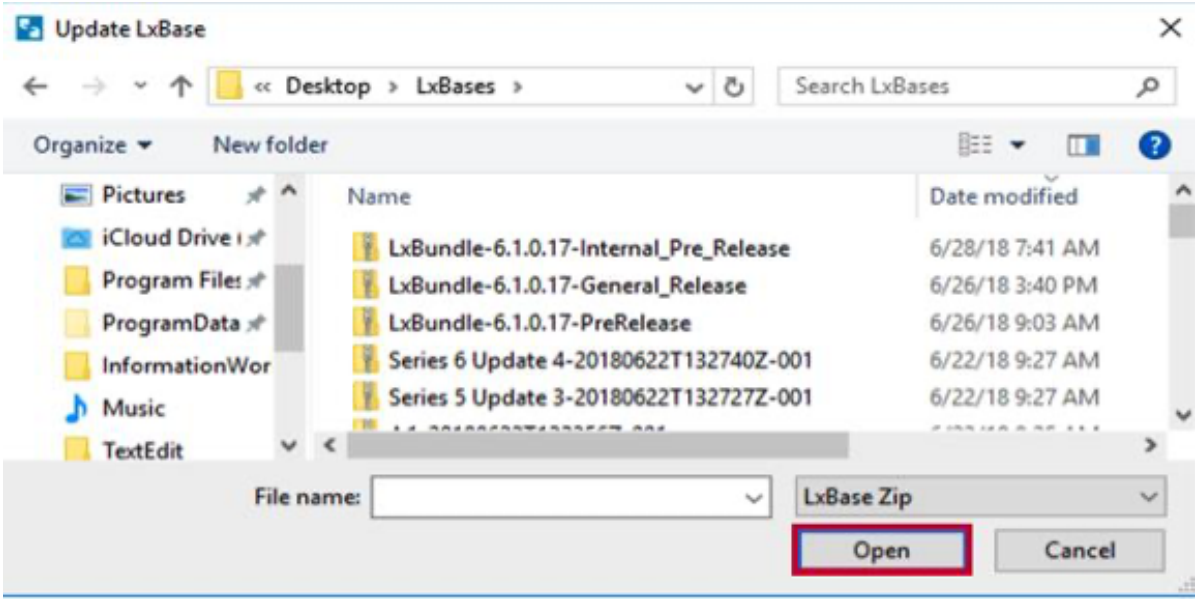
To update your LxBase, you use the TextChart Text Analytics Configuration application to provide i2 TextChart with the location of a ZIP file that contains the new information.

1. Open the configuration application and select the **Assets** tab:



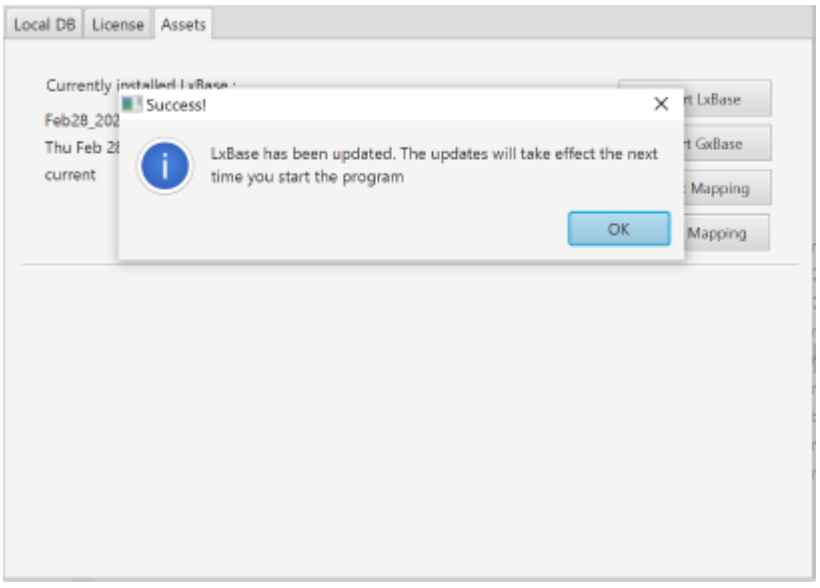
2. Click **Import LxBase**, navigate to the new ZIP file, and click **Open**.

© N. Harris Computer Corporation



i2 TextChart automatically updates your LxBase with information from the file.

3. When the following dialog appears, click **OK**:



The changes and updates will take effect when you restart Analyst's Notebook.

Note: i2 TextChart backs up the previous version of your LxBase in `C:\ProgramData\TextChart`, in a directory named `LxBase-backup-<TIMESTAMP>`. To restore a backup, delete the `LxBase` directory, and then rename the backup to `LxBase`.

© N. Harris Computer Corporation

Upgrading i2 TextChart Standard

Upgrading i2 TextChart Standard requires a complete uninstallation of the previous version before installing the new release. This procedure ensures a clean installation and prevents potential conflicts between versions.

About this task

The upgrade process consists of the following stages:

1. Back up your existing database configuration
2. Uninstall the current version of i2 TextChart Standard
3. Install the new version using the standard installation procedure

Prerequisites

Before beginning the upgrade:

- Ensure you have administrator privileges on the workstation
- Verify i2 TextChart Standard and i2 Analyst's Notebook are closed
- Obtain the installation package for the new version
- Close all related applications before starting the upgrade process

Procedure

Step 1: Prepare for the upgrade

1. Close i2 Analyst's Notebook.
2. If you are using a local TextChart database, create a backup:

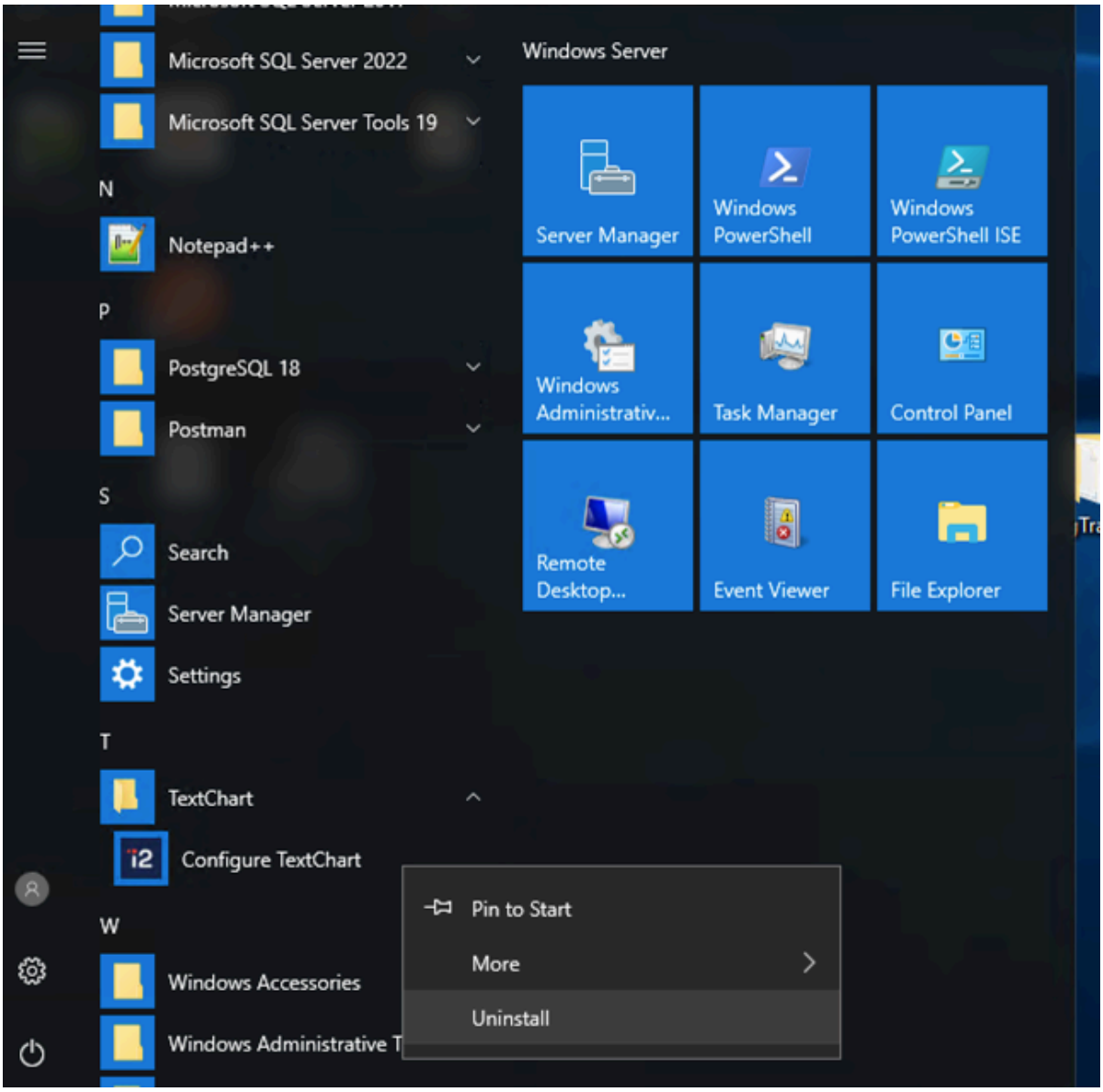
   a. **For SQL Server databases**: Use SQL Server Management Studio or other database administration tools to create a full backup of your TextChart database.

   b. **For embedded databases**: Navigate to the database directory at `C:\ProgramData\TextChart\databases\Rosoka` (or `C:\ProgramData\Rosoka\TextAnalytics\databases\Rosoka` for older versions) and create a copy of the entire folder structure.

Step 2: Uninstall the current version

1. Uninstall i2 TextChart Standard using one of the following methods:
   - From the Windows Start menu, locate and run the TextChart uninstaller
   - Navigate to `C:\Program Files\TextChart\` and execute `Uninstall.exe`
   - Use Windows **Add or Remove Programs** from the Control Panel

© N. Harris Computer Corporation



2. Follow the uninstaller prompts and wait for the process to complete before proceeding.

Step 3: Install the new version

After successful removal of the previous version, install the new version by following the procedures documented in Installing i2 TextChart Standard.

What's next

After completing the installation, restore your database content from the backup created in Step 1, if required. Test the installation by processing a sample document to verify that all components are functioning correctly.

© N. Harris Computer Corporation

Consider reviewing the release notes for the new version to understand any new features or configuration changes that may benefit your deployment.

Other configuration

i2 TextChart provides a few ways to modify its behavior so that it works better in your environment. For example, you can adjust the amount of memory available to the extraction process, and change the way TextChart authenticates with the database.

Memory configuration

If you experience poor performance, especially during processing of large files, you can change the size of the JVM heap by editing the configuration file at `C:\Program Files\TextChart \RosokaServiceCore.l4j.ini`.

To adjust memory allocation, which defaults to 4 GB, change:

```
-Xms4g
-Xmx4g
```

to:

```
-Xms<GB_OF_RAM_TO_ALLOCATE>g
-Xmx<GB_OF_RAM_TO_ALLOCATE>g
```

For example, to set it to 8 GB, the two lines would be:

```
-Xms8g
-Xmx8g
```

Database configuration (SQL Server)

TextChart authenticates with SQL Server through database user credentials rather than operating system credentials. You need to ensure that your installation of SQL Server is configured appropriately.

Setting mixed authentication

1. Open SQL Server Management Studio.
2. From the Authentication drop-down list, select **Windows Authentication** .
3. Click **Connect**.
4. Right-click the SQL instance and choose **Properties** to open the Server Properties window.
5. In the **Select a page** list, click **Security**.
6. In the **Server authentication** section, choose **SQL Server and Windows Authentication mode**.
7. Click **OK** to close the Server Properties window.
8. Right-click the SQL instance and choose **Restart**.
9. Click **Yes** to restart the SQL server.

Enabling the 'sa' user

1. Open SQL Server Management Studio.
2. From the Authentication drop-down list, select **Windows Authentication**.
3. Click **Connect**.
4. In the Object Explorer, expand **Security** and then expand **Logins**.
5. Right-click **sa**, and click **Properties**.

© N. Harris Computer Corporation

6. In the **Select a page** list, click **Status**, and then change the **Login** setting to **Enabled**.

   Note: You might need to set a password for this account as well.
7. Click **OK** to close the Login Properties window.
8. Press *F5* to refresh the information in the Object Explorer.

Changing TCP and port settings

1. From the **Start** menu, right-click **My Computer**, and then select **Manage**.
2. In the Computer Management window, expand **Services and Applications > SQL Server Configuration Manager > SQL Server Network Configuration**.
3. Select the instance for your installation. By default, this is **Protocols for SQLEXPRESS**.
4. Change the **Named Pipes** protocol option to **Enabled**. You do not need to change the default **Pipe Name** setting.
5. Change the TCP/IP protocol option to **Enabled**, and then open its **Properties** dialog.
6. In the **IP Addresses** tab, scroll down to **IPAll** at the bottom of the list.
7. Set **TCP Port** to 1433, and remove any value from **TCP Dynamic Ports**.
8. Click **Apply**. You will be prompted to restart SQL Server.
9. Restart the server for the changes to take effect.

Importing and processing documents

The i2 TextChart plug-in for Analyst's Notebook processes documents to extract information about the things they describe and the connections between them. You can add the extracted information to your charts in the form of entities and links.

Entity processing

Entities are the important, named items, such as the people, places, and events that a document describes. i2 TextChart uses the linguistic context of the document to determine what words or phrases are extracted as entities.

After processing, you can modify entity extraction results and apply your own knowledge before sending results to a chart.

Note: The LxBase documentation includes a list of all the entity types that TextChart understands, with their corresponding definitions.

Link processing

Links are relationships between two entities that i2 TextChart establishes through linguistic context. In extraction results, links appear in the form `<Entity>To<Entity>`. For example, `PersonToPerson` means that there is a link between one Person entity and another Person entity.

The predicate (that is, the reason why i2 TextChart identified the link) is stored with the link. If the predicate is "interviewed", then the link indicates that a Person interviewed another Person.

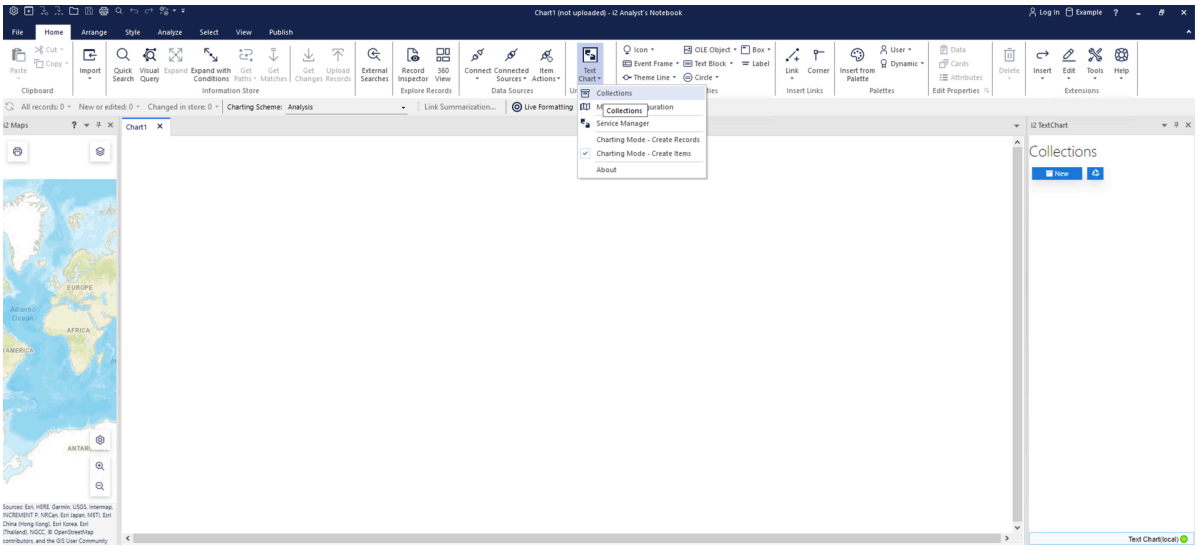
Note: The LxBase documentation includes a list of all the predicate types that TextChart understands.

Importing documents

To use i2 TextChart, the first step is to import the documents that you want to analyze.

© N. Harris Computer Corporation

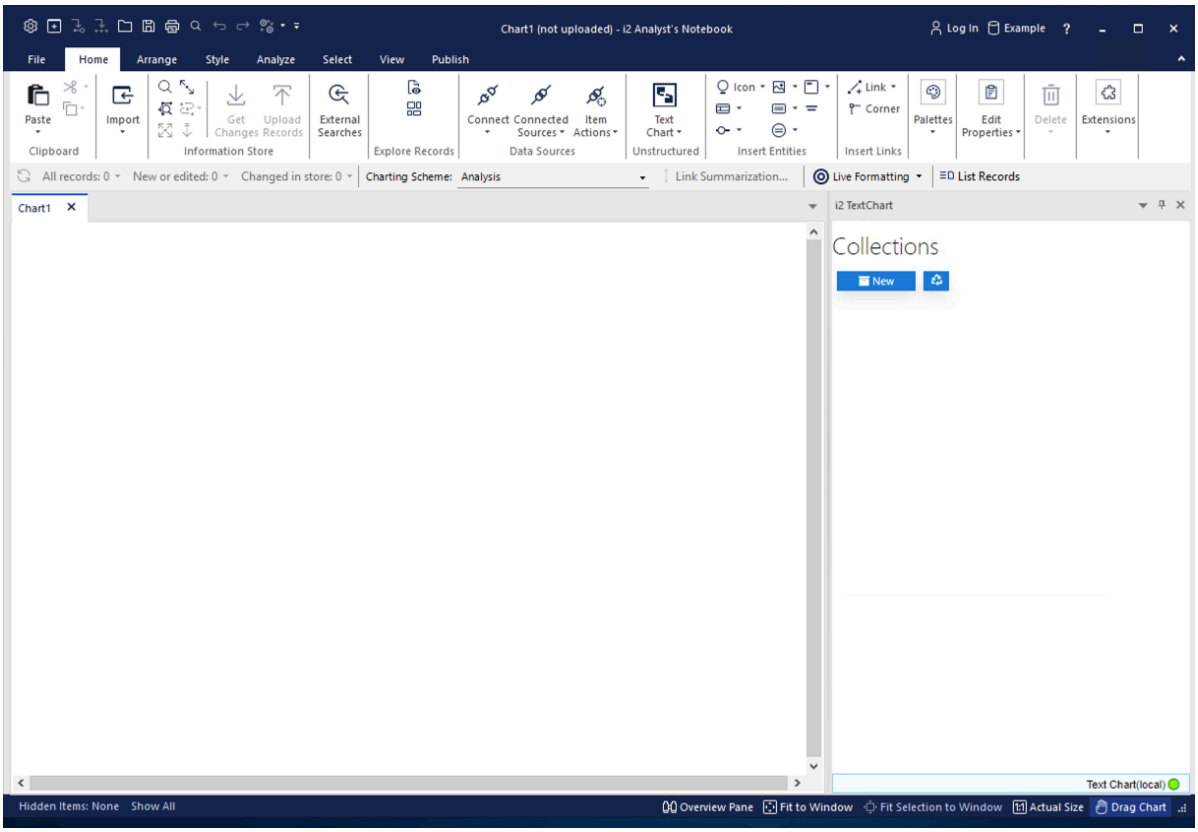
1. In Analyst's Notebook, select the **Text Analytics** menu in the **Home** tab, and then click **Collections**.



Note: The Mapping Configuration option is available only in the standard edition of TextChart. In TextChart Premium, mapping takes place on the server.

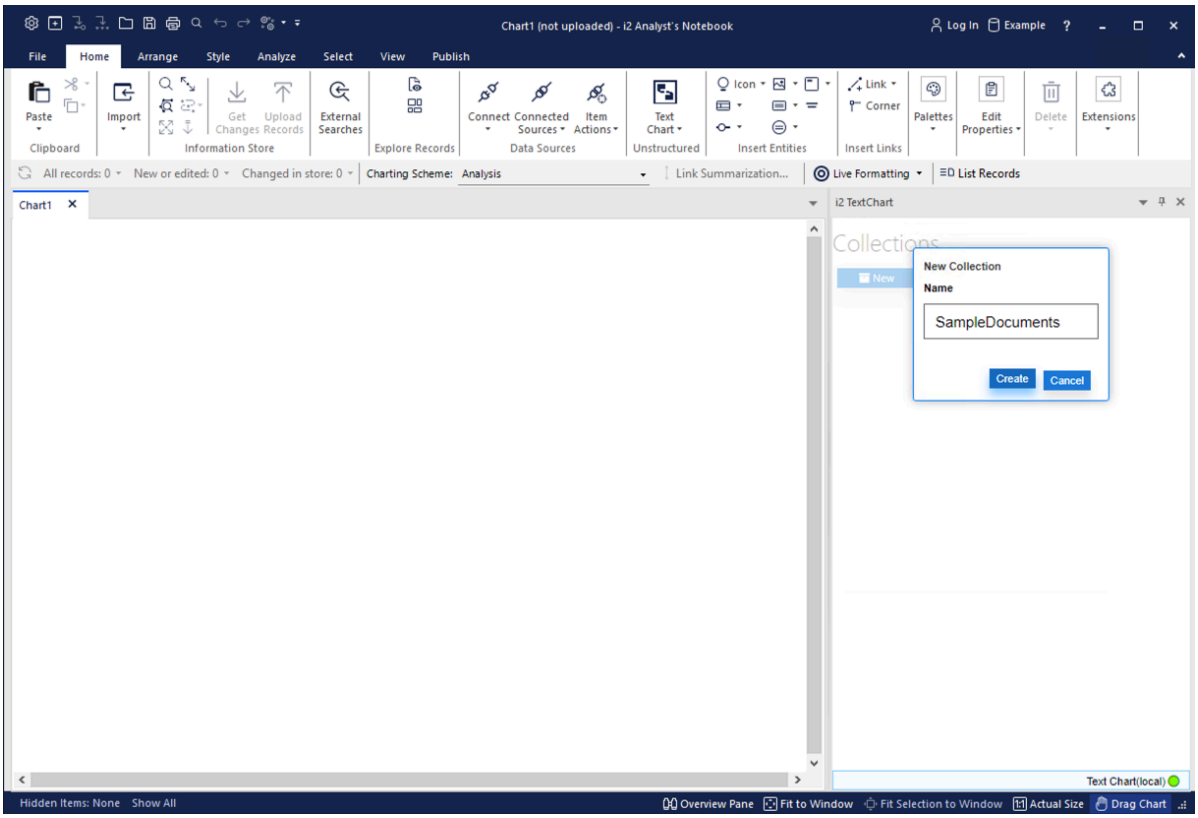
2. i2 TextChart stores documents in *collections*. A collection contains one or more documents that you choose to be in the same group for processing. You can start a new collection for the documents that you're importing, or add them to an existing one.

© N. Harris Computer Corporation



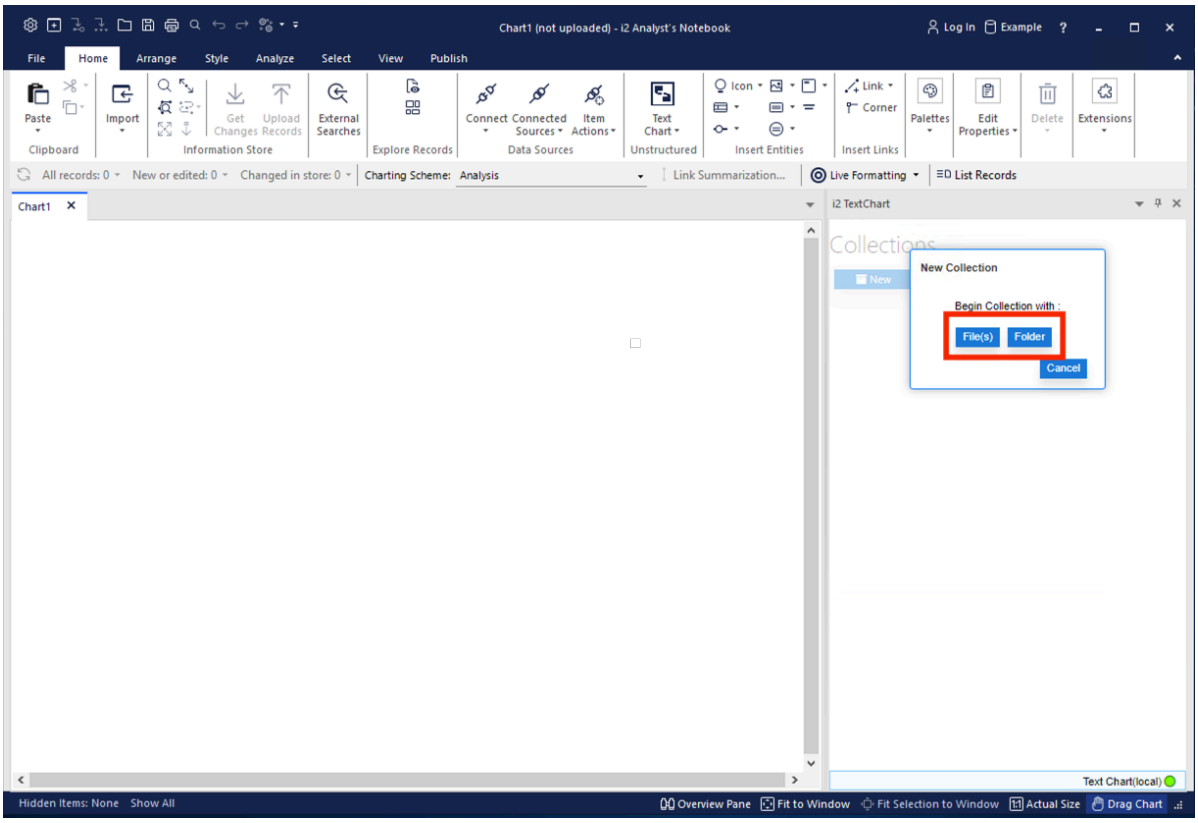
3. If you clicked **New**, provide a name for the new collection and click **Create**.

© N. Harris Computer Corporation



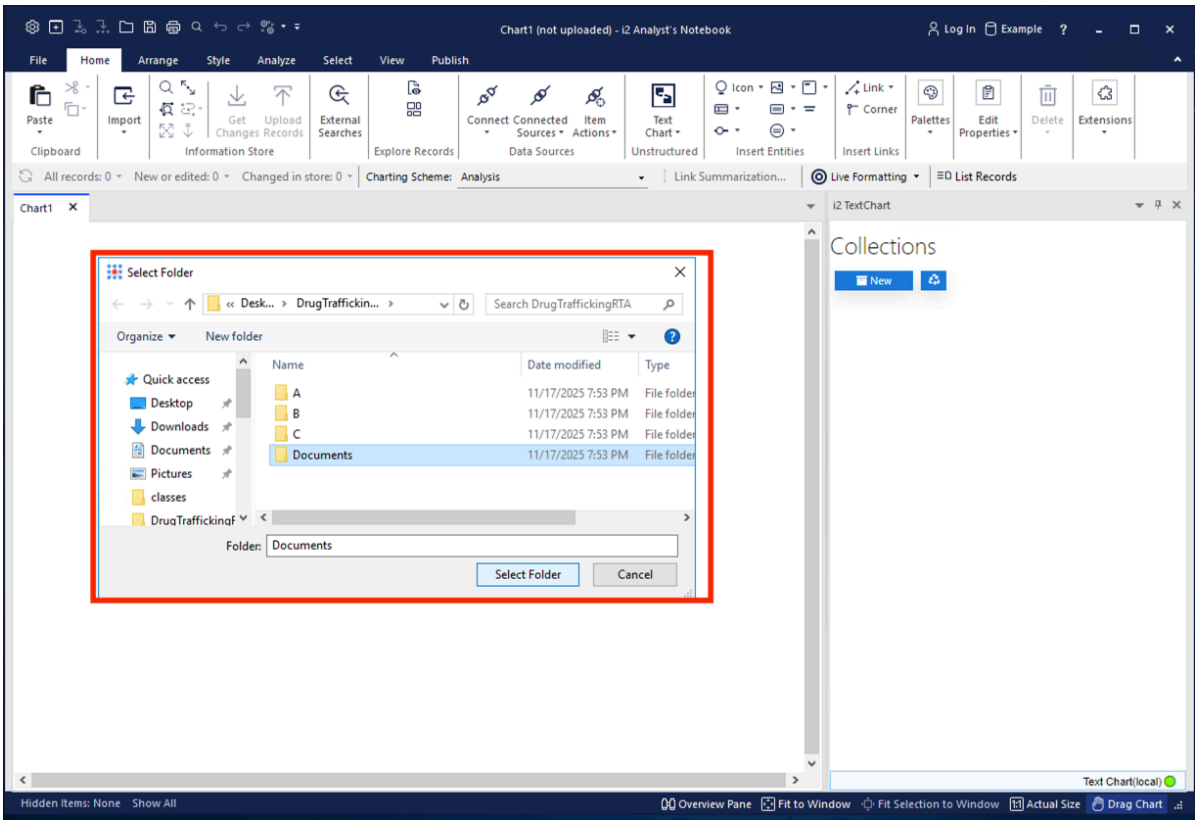
And then specify whether to populate the new collection with files or a folder.

© N. Harris Computer Corporation



4. Finally, select the files (or the folder that contains the files) that you want to import.

© N. Harris Computer Corporation



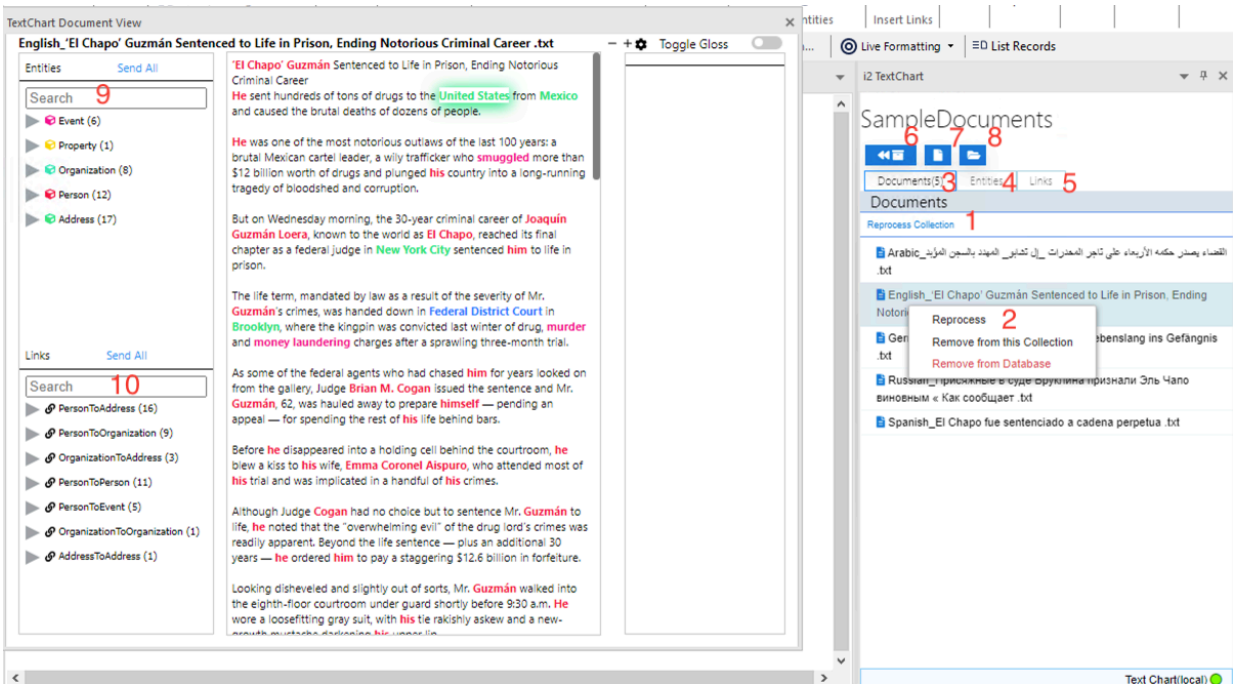
Processing documents

After you import documents into a collection, you can instruct TextChart to process them all at the same time, or to process each document separately.

Note: When you process documents for the first time, it can take a few seconds for the engine to load. Also, larger documents generally take longer to process than shorter ones. TextChart displays a spinning cursor while long-running operations take place.

The documents in a particular collection appear as a list in Analyst's Notebook, from which you can perform a number of tasks. Clicking a single document in the list processes that document and opens it in the Document View. Some of the other tasks are described below.

© N. Harris Computer Corporation



1. **Process Collection**

   When you click **Process Collection**, TextChart processes all the documents in the list that it has not already processed. The progress bar at the top of the pane shows the status of the operation. After each document is processed, its page icon turns blue.

   Canceling the operation stops processing for the current and subsequent documents in the collection. You can also ask TextChart to reprocess *all* documents in the collection, regardless of whether they've been processed before.

2. **Reprocess** / **Remove from this Collection** / **Remove from Database**

   If you right-click a single document in the list, the pop-up menu presents commands to reprocess the document, or to remove it from the collection, or to remove its processing results from the database.

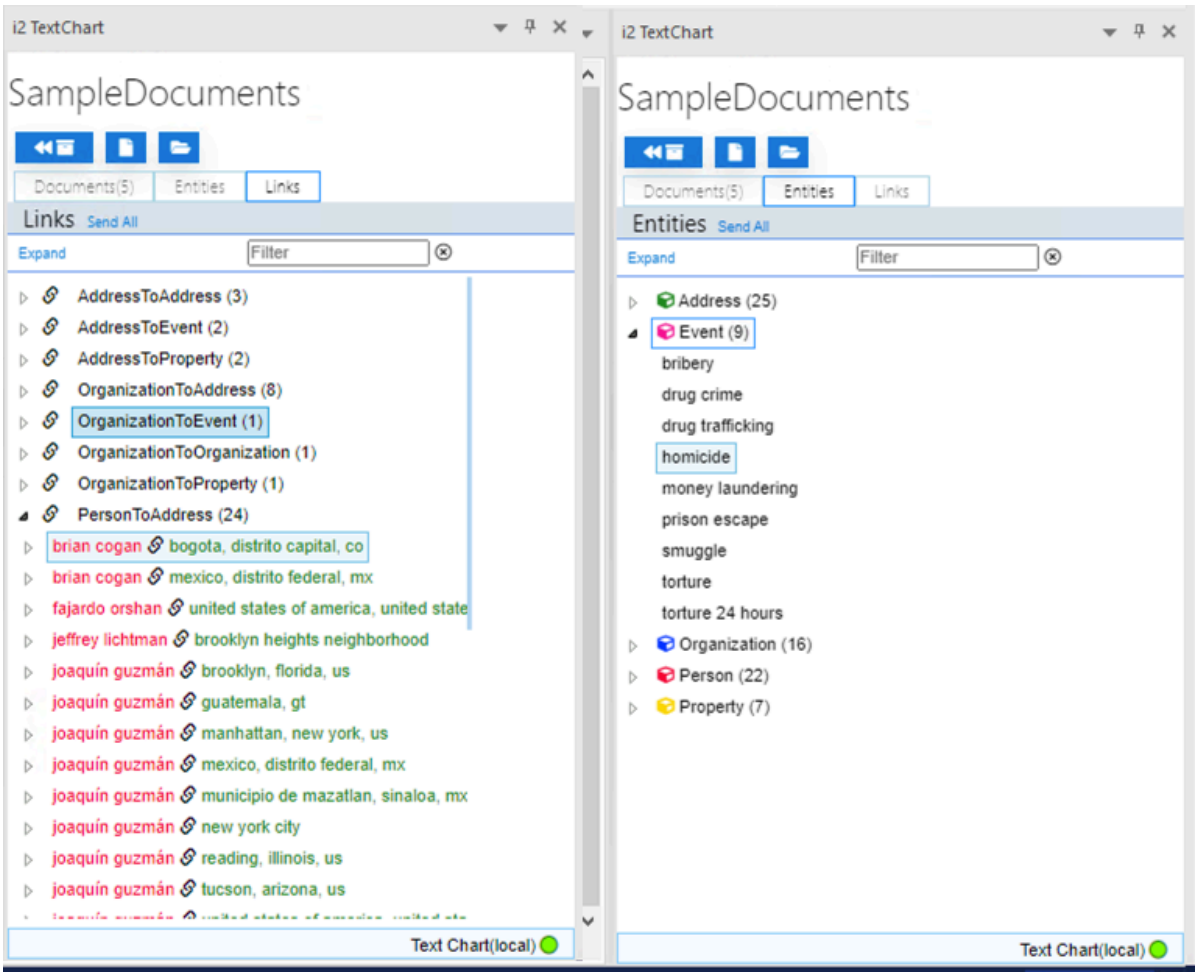
3. **Documents**

   The Documents tab displays a count of the documents in the collection.

4. **Entities**

   The Entities tab contains an entity tree, which lists the TextChart types of all the found entities in processed documents, as well as the count of each type of entity.

© N. Harris Computer Corporation



5. **Links**

   The Links tab contains a link tree that lists the TextChart types of all the found links in processed documents, as well as the count of each type of link.

   You can expand the link types in the tree view to see lists of the individual links and the predicate types that TextChart identified during processing.

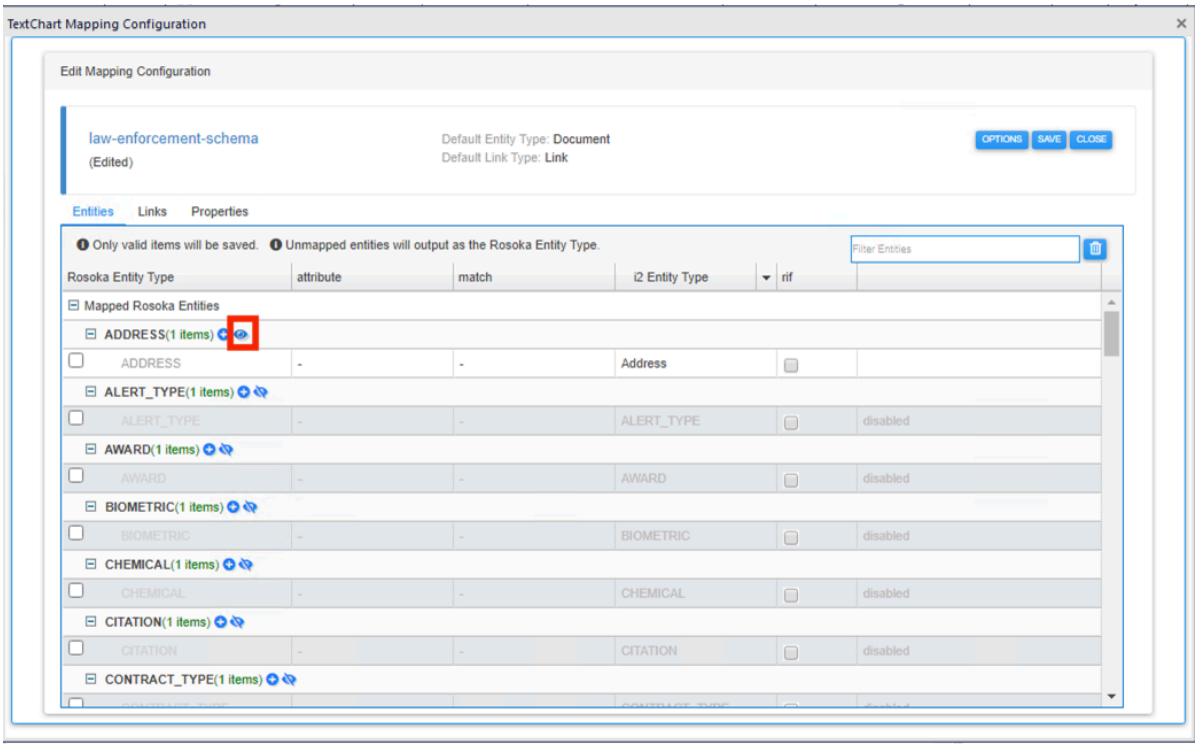
6. Click **Back to Collections** to return to the list of collections.
7. Click **Add More Files** to add more files to the current collection.
8. Click **Add Directory to Collection** to add all the documents in a particular directory to the current collection.

Ignoring TextChart entity, link, or attribute types

By default, TextChart attempts to find entities and links of all the types that it understands, and values for all the attributes they can contain. If there are some types that you don't want TextChart to extract, you can disable them through the Mapping Configuration.

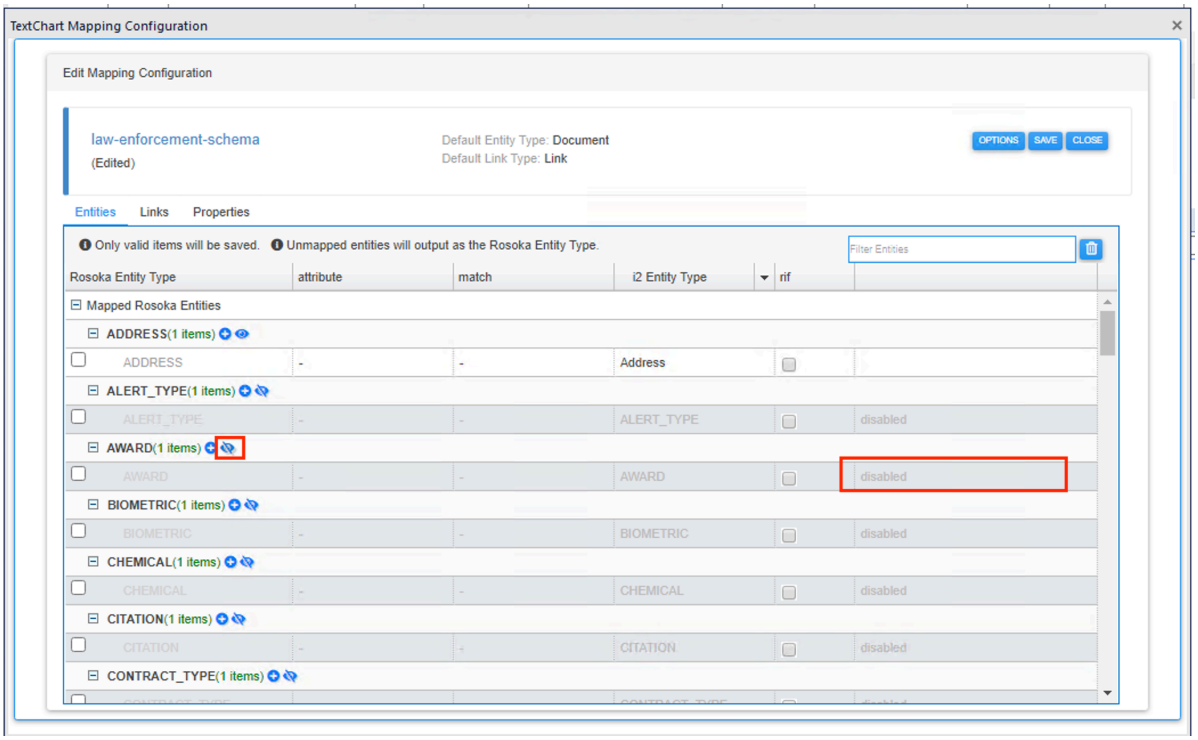
© N. Harris Computer Corporation

1. From the Analyst's Notebook ribbon, select **Text Analytics > Mapping Configuration** to open the Mapping Configuration window.
2. Click **Edit** to display the configuration settings for the active mapping. The left column of the **Entities**, **Links**, and **Properties** tabs contains a list of TextChart types.



3. Click the "eyeball" icon next to the name of any type that TextChart should not extract from the documents it processes.

© N. Harris Computer Corporation



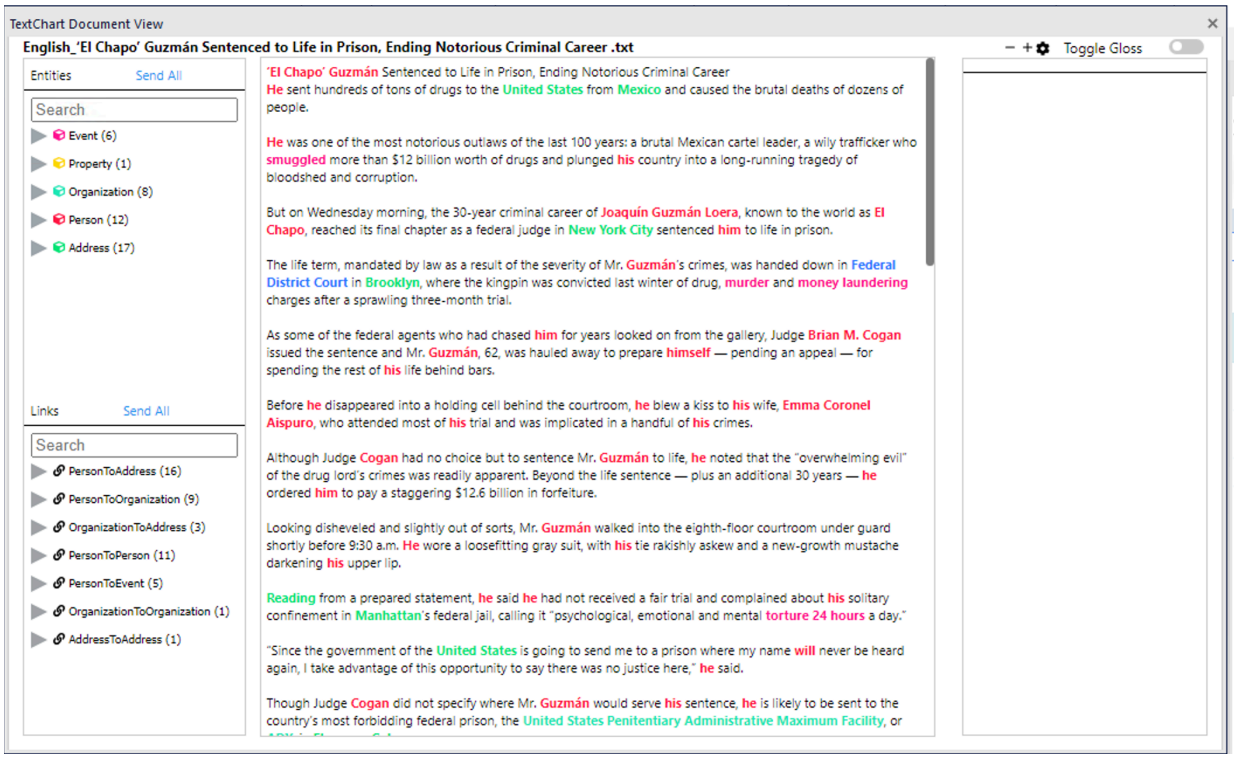
4. If you processed some documents before changing these settings, clear those results and reprocess the documents.

Viewing documents

After TextChart processes a document (or the first of several documents), it opens the Document View window, which presents an enriched view of the processed text.

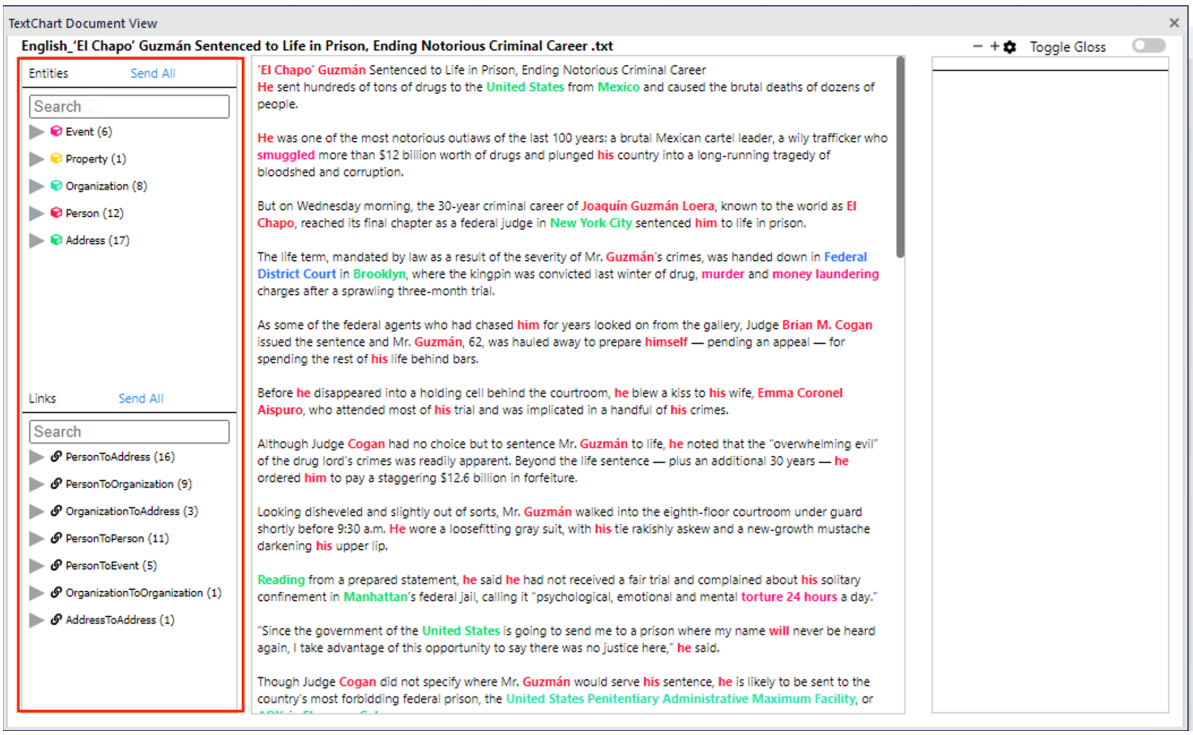
The Document View highlights all of the found entities in context, using different colors for different types. In this view, you can review and curate the entities that TextChart identified.

© N. Harris Computer Corporation



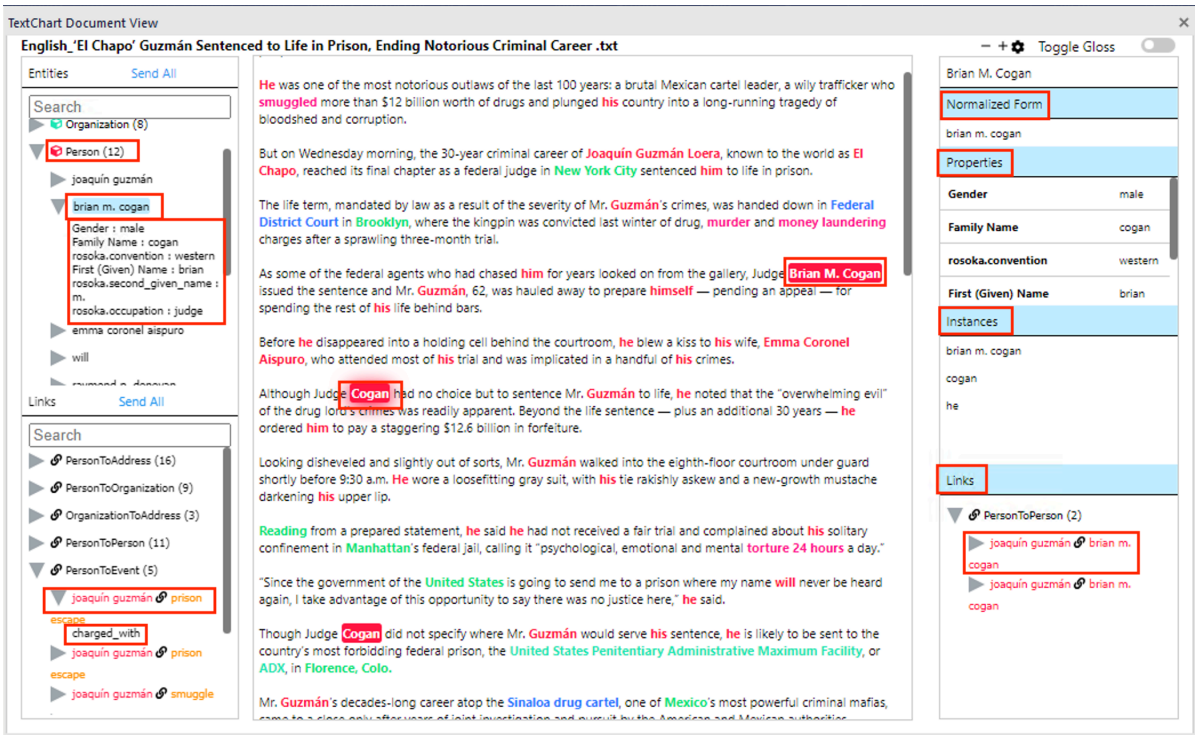
1. The button at the upper left of the window opens entity and link trees for the current document.

© N. Harris Computer Corporation



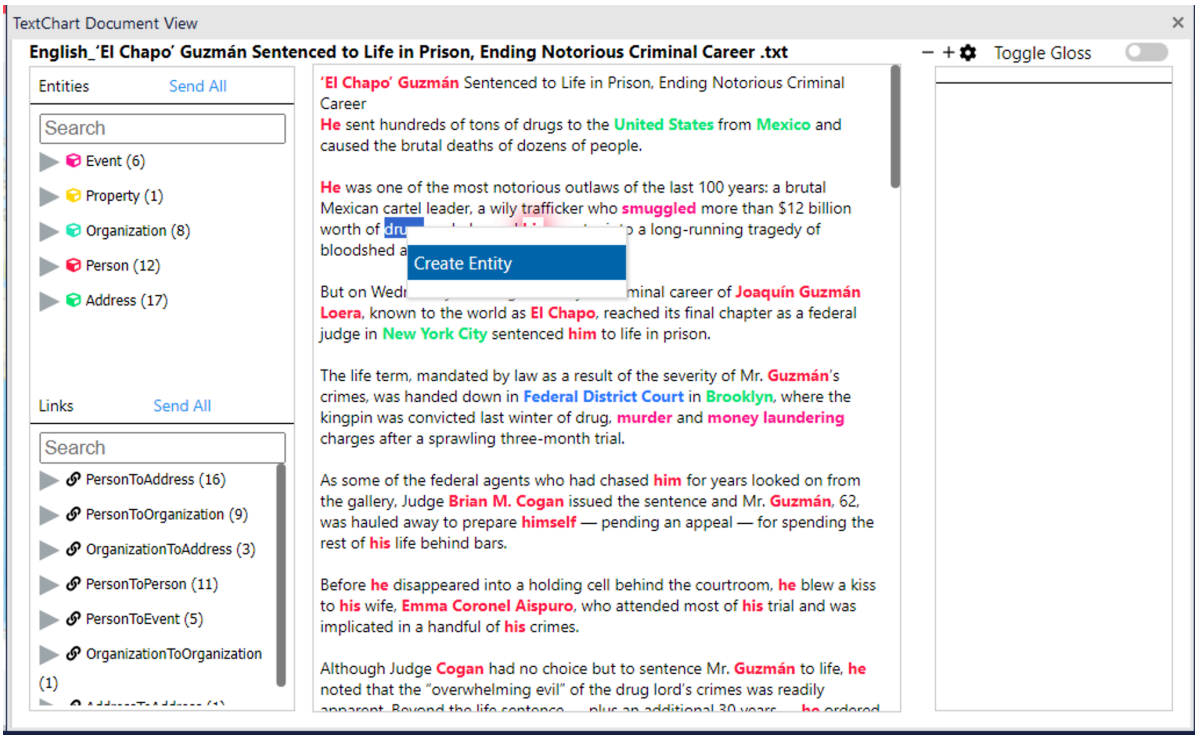
The trees behave initially like the ones in the list for all the documents in a collection. However, when you select entities and links (or groups of them), TextChart highlights them in the open document and displays a second panel that contains more information.

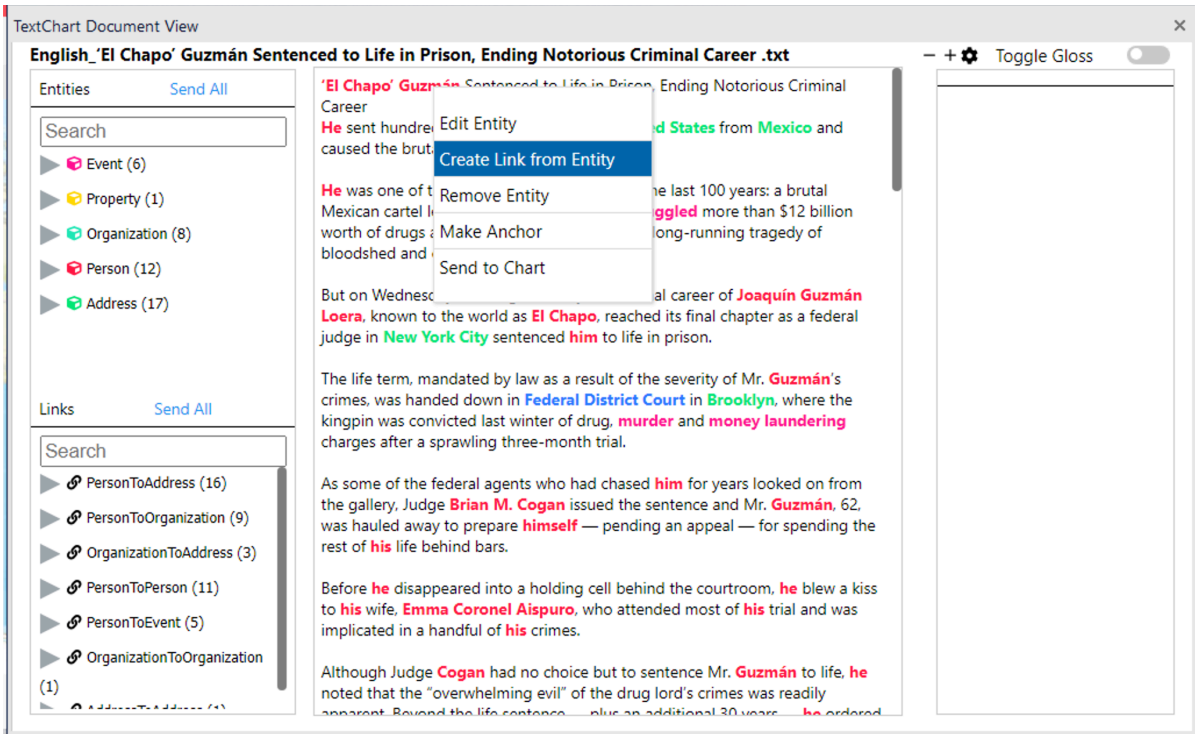
© N. Harris Computer Corporation



2. Use the **Tag** drop-down to tag a new entity or link.

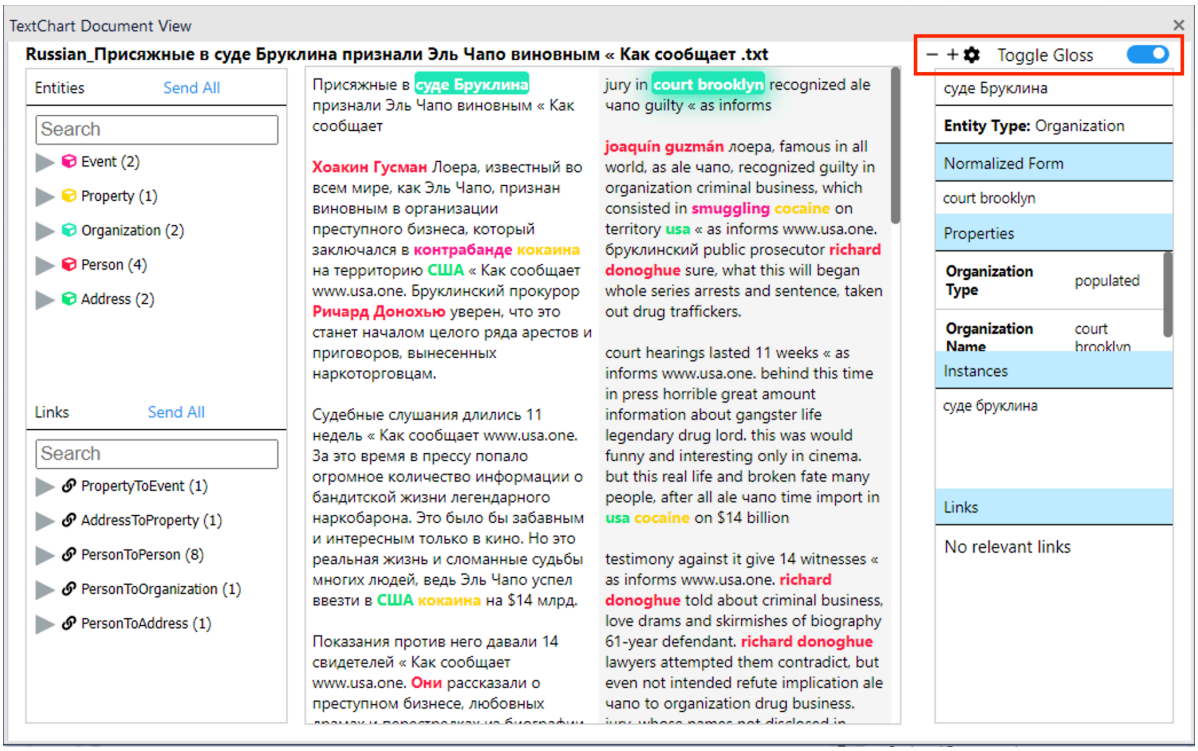
© N. Harris Computer Corporation





© N. Harris Computer Corporation

3. Through the **Options** button, you can:
   a. Toggle back and forth between an English version of the document, and the text in its original language.
   b. Remove whitespace from the document as displayed
   c. Change the font size of the text in the document.



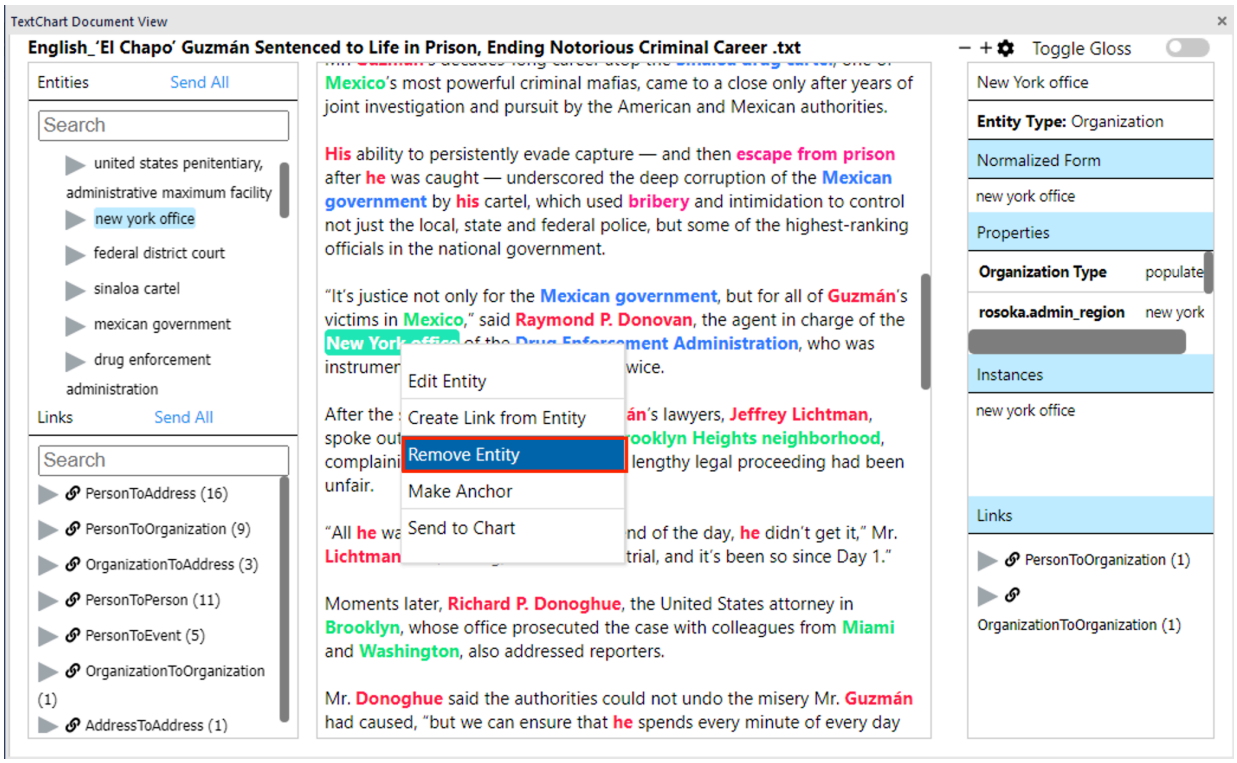
Curating extraction results

After i2 TextChart presents the results of processing a document in the Document View, you can review its output and make changes to the results. TextChart supports removing and modifying the records it finds, as well as identifying new records of your own.

Removing results

If TextChart identifies an entity or a link that you don't want to appear in the results, you can remove it from the Document View. For example, the following view contains a contrived entity extraction result for a "New York office", which has the Organization entity type.

© N. Harris Computer Corporation

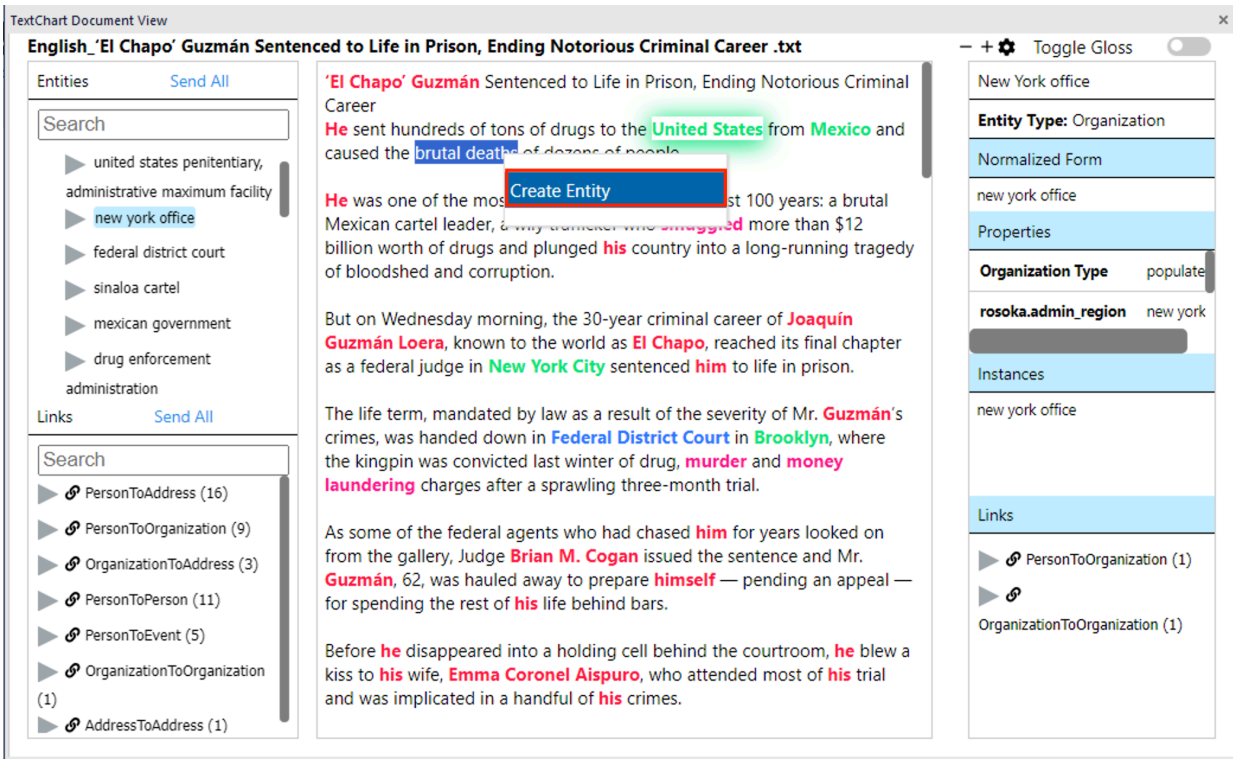


When you select the text in the Document View, TextChart displays information about the extracted entity. To remove this result, click **Remove Entity**.

Adding entities

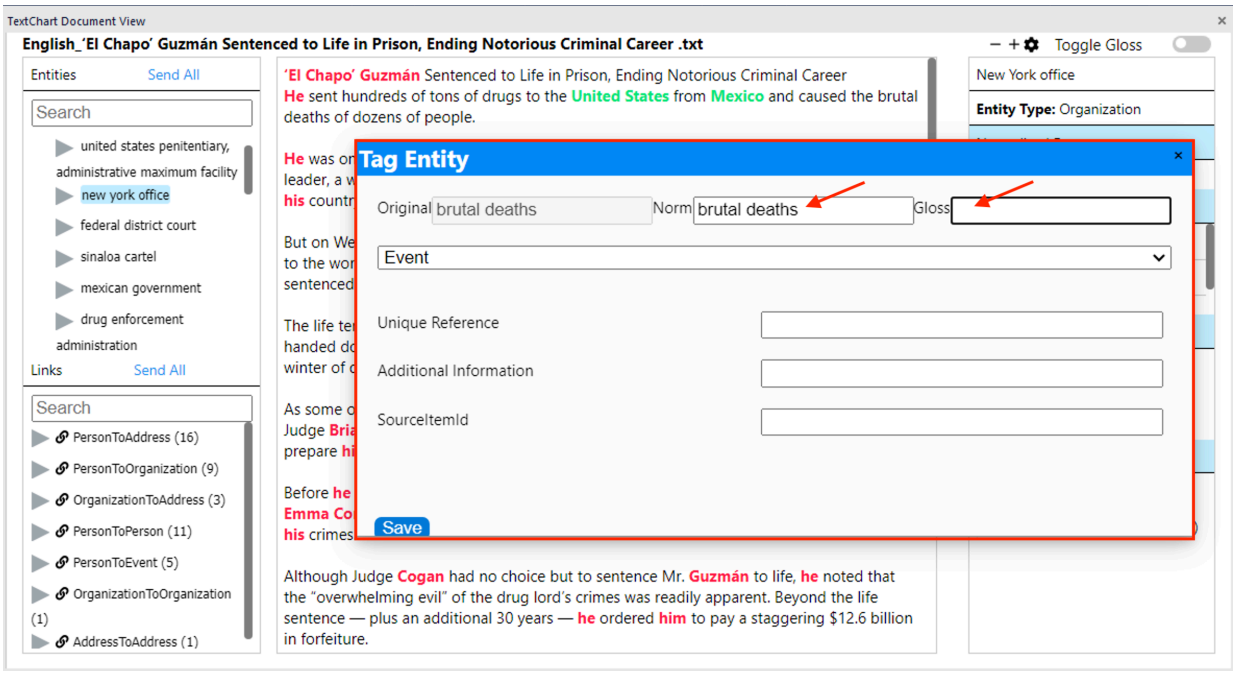
If TextChart failed to identify an entity that you *do* want to appear in the results, you can highlight the text to be extracted, and then select **Create Entity**.

© N. Harris Computer Corporation



TextChart displays the Tag Entity Tool for you to provide information about the entity, and populates the **Original** field with the text that you highlighted.

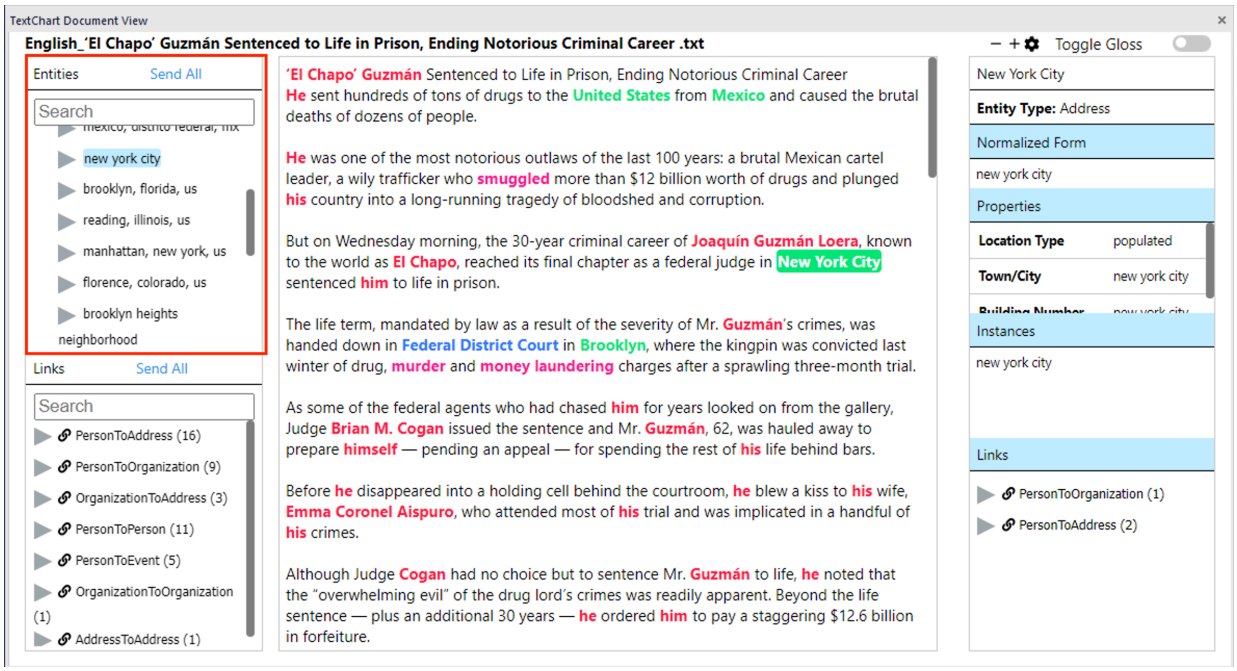
© N. Harris Computer Corporation



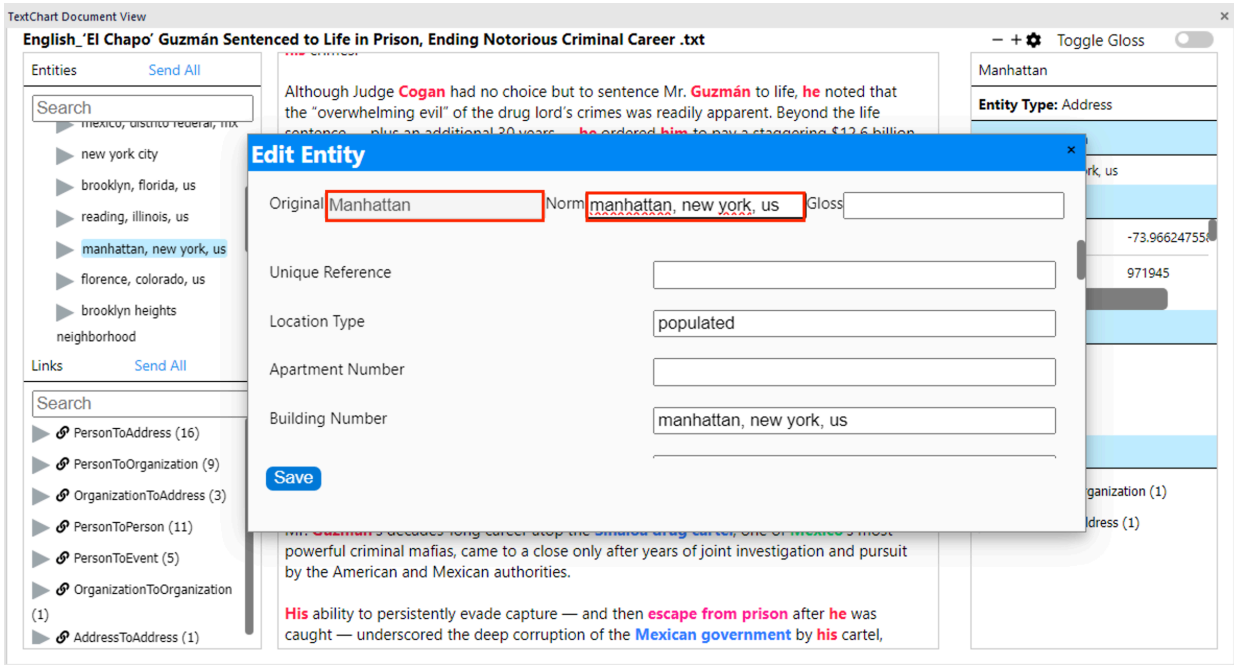
TextChart also populates the **Norm** field with the same text, but here you have the option of changing the text to normalize it when a document uses different terminology to refer the same piece of information.

For example, in the image below, "new york city" and "manhattan, new york, us" have been identified as separate places.

© N. Harris Computer Corporation



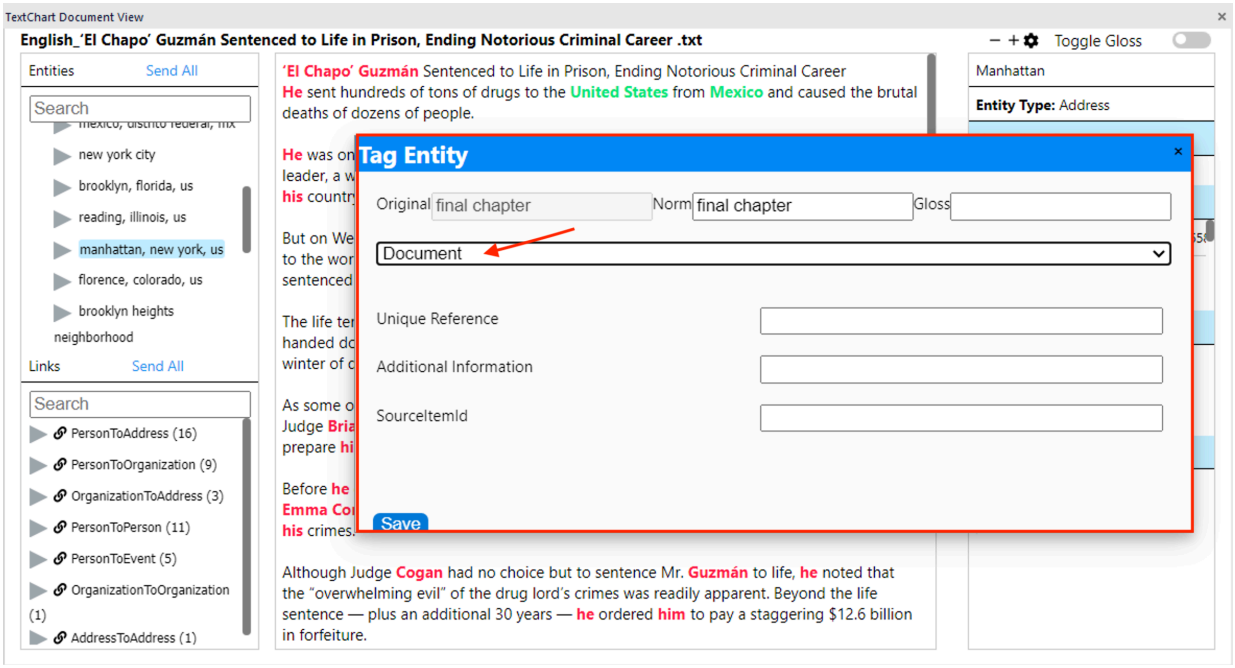
If you decide that TextChart should treat these instances as the same place, you can edit the **Norm** field of one so that it matches the other.



© N. Harris Computer Corporation

If you're processing documents in a language other than English, and you come across an important term that TextChart does not understand, you can add an English translation of that term to the **Gloss** field.

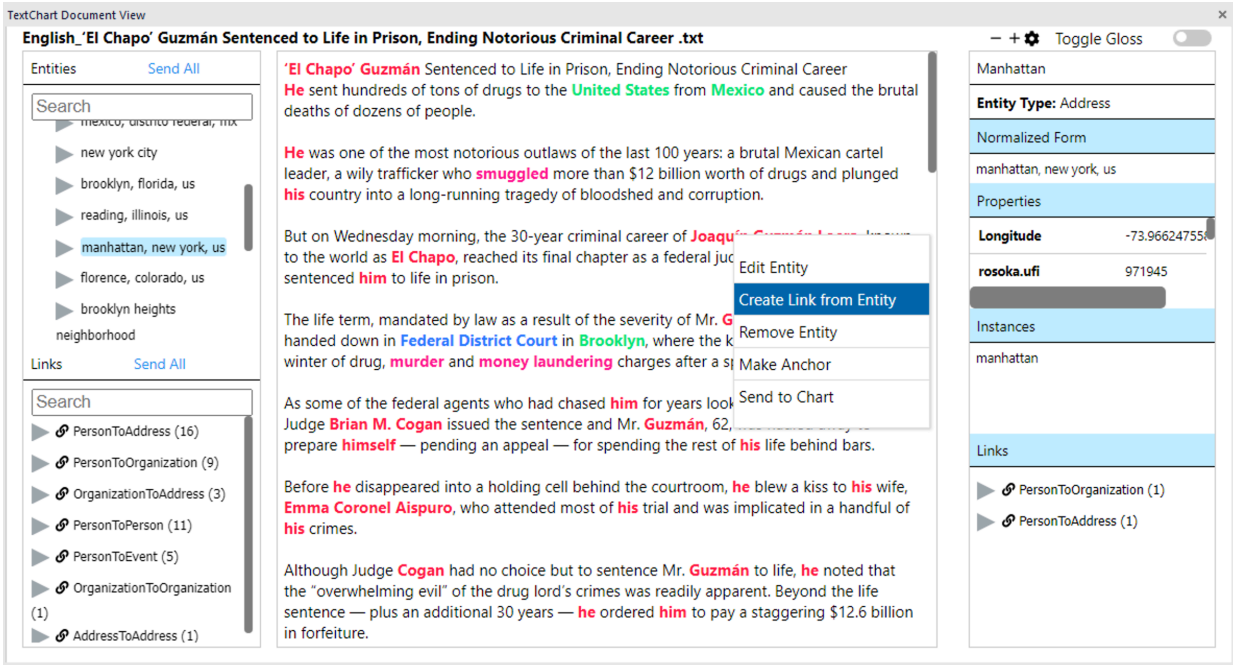
Finally, you must choose a TextChart type for your new entity. Different entity types have different attributes, and you can fill in additional information as you see fit.



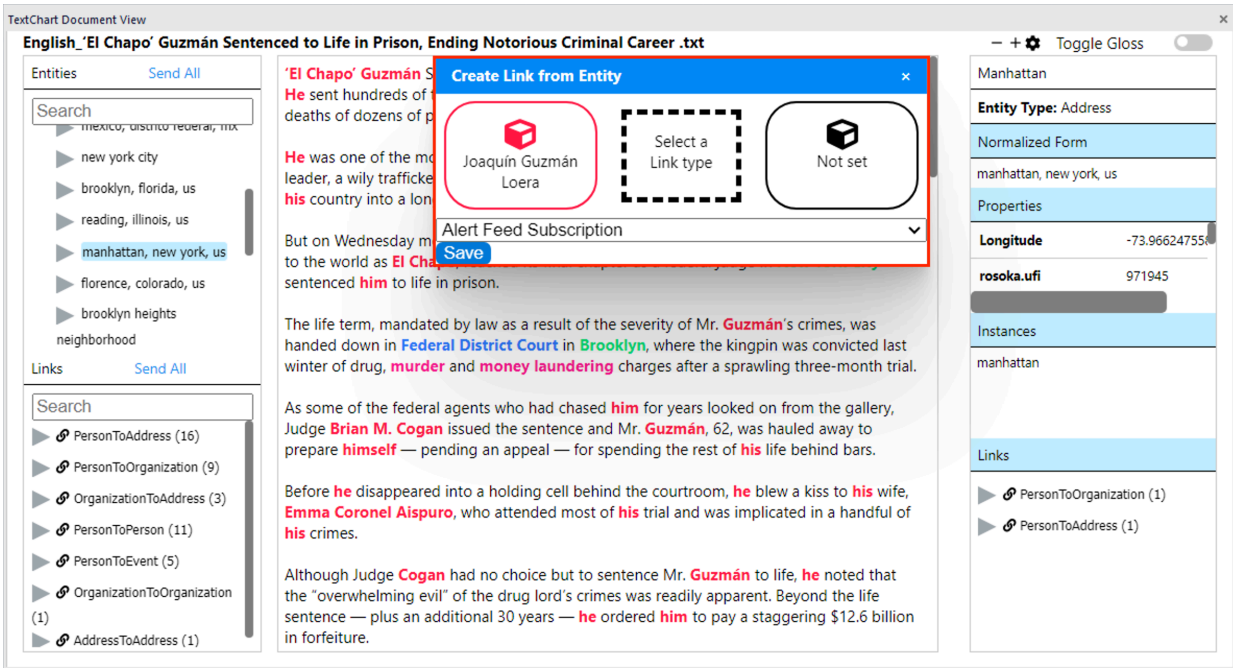
Adding links

Adding a link that TextChart failed to identify is similar to adding an entity in the same circumstances. You can highlight, right-click an extracted entity and select **Create Link from Entity** to use it as the "From" end of a new link.

© N. Harris Computer Corporation

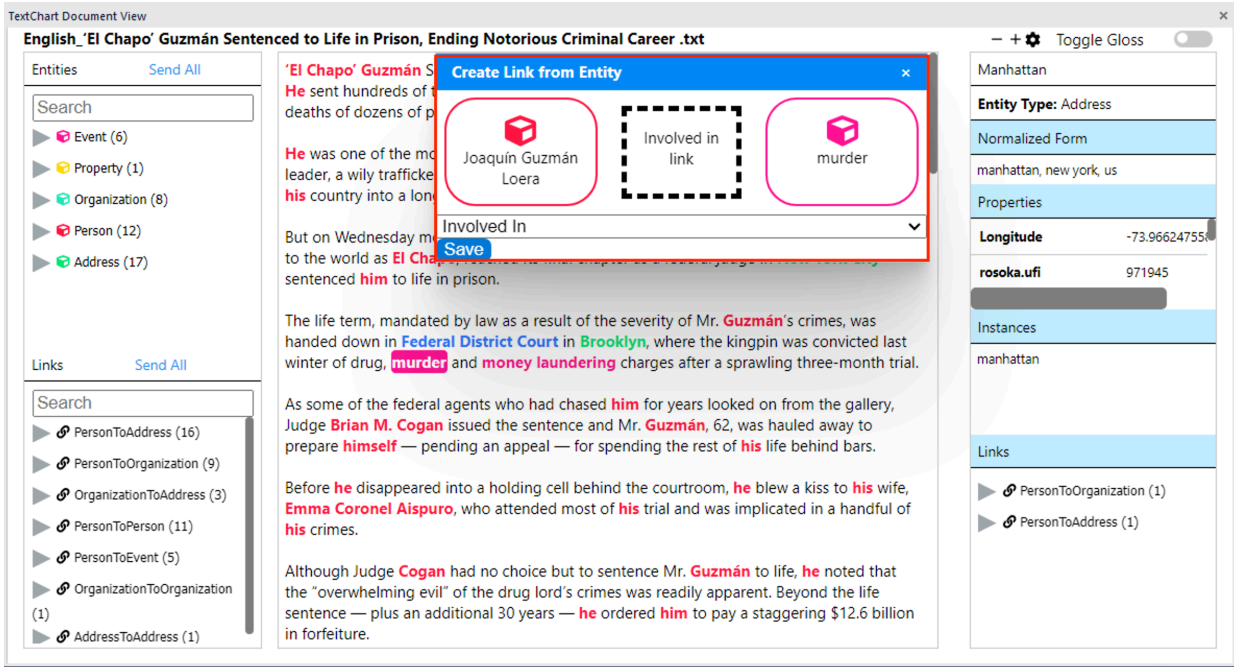


The Tag Link Tool starts by prompting you to select the entity at the "From" end.

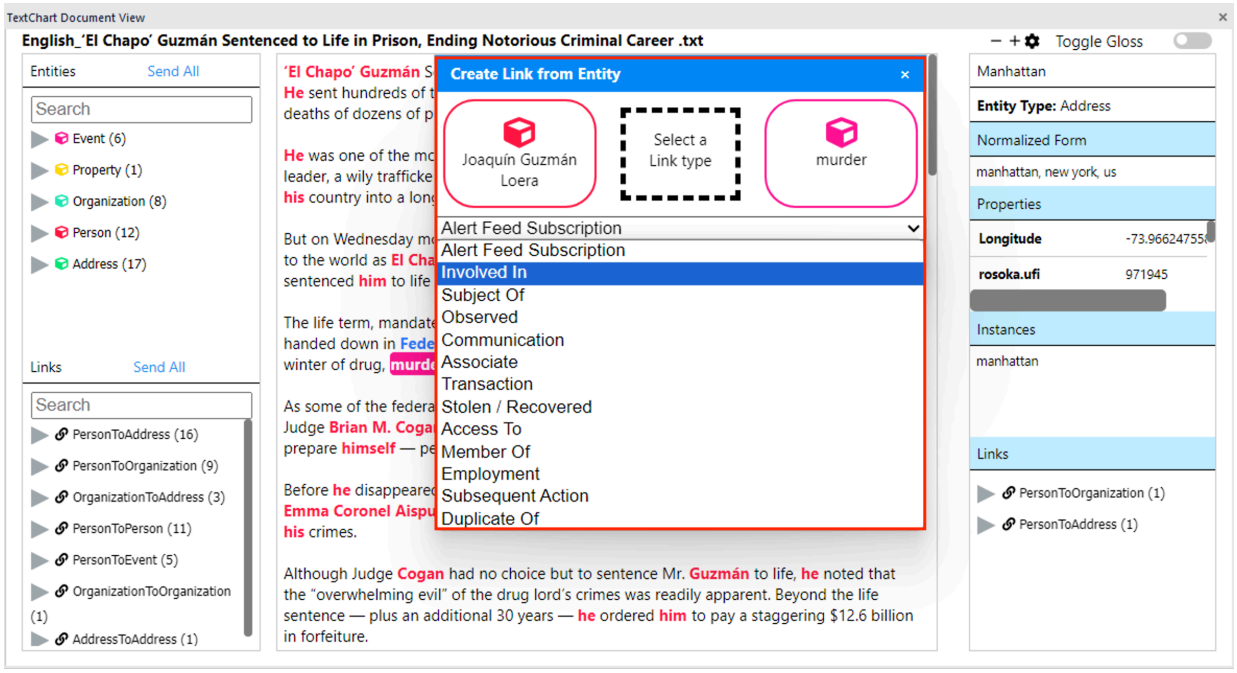


Then, you have to select the entity at the "To" end.

© N. Harris Computer Corporation

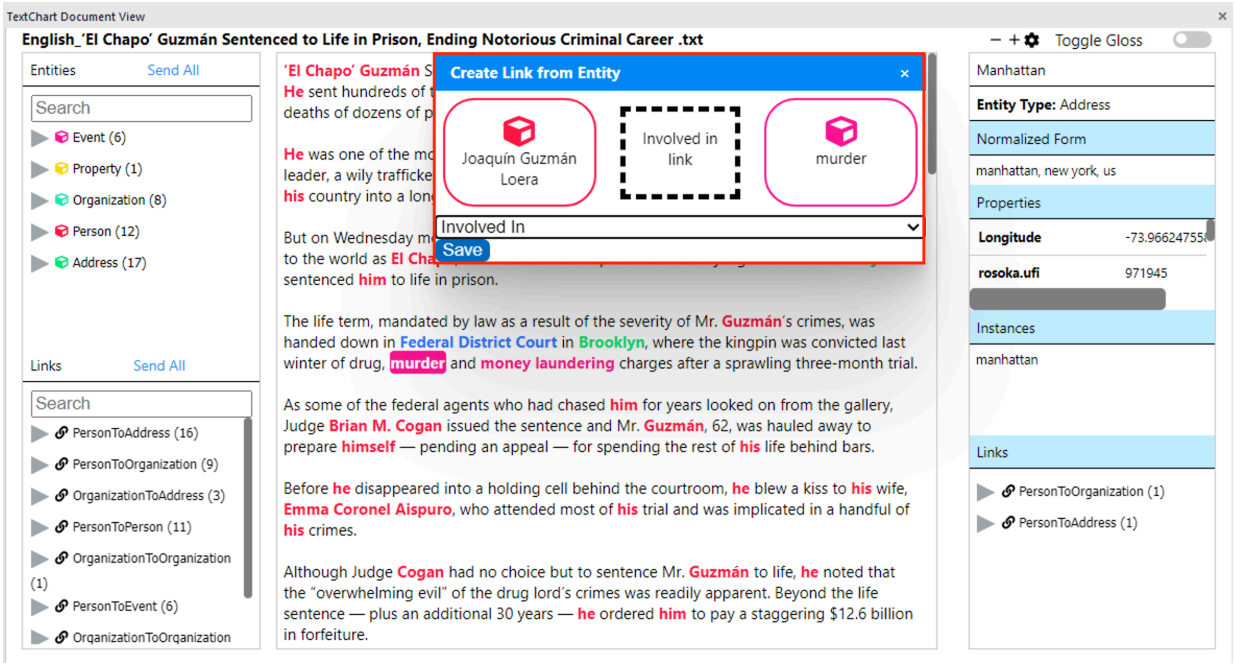


After you select both ends, the Tag Link Tool displays a drop-down that allows you to specify the TextChart type of the link to be created.



When all the information in the Tag Link Tool is complete, click **Save** to add the new link to the results.

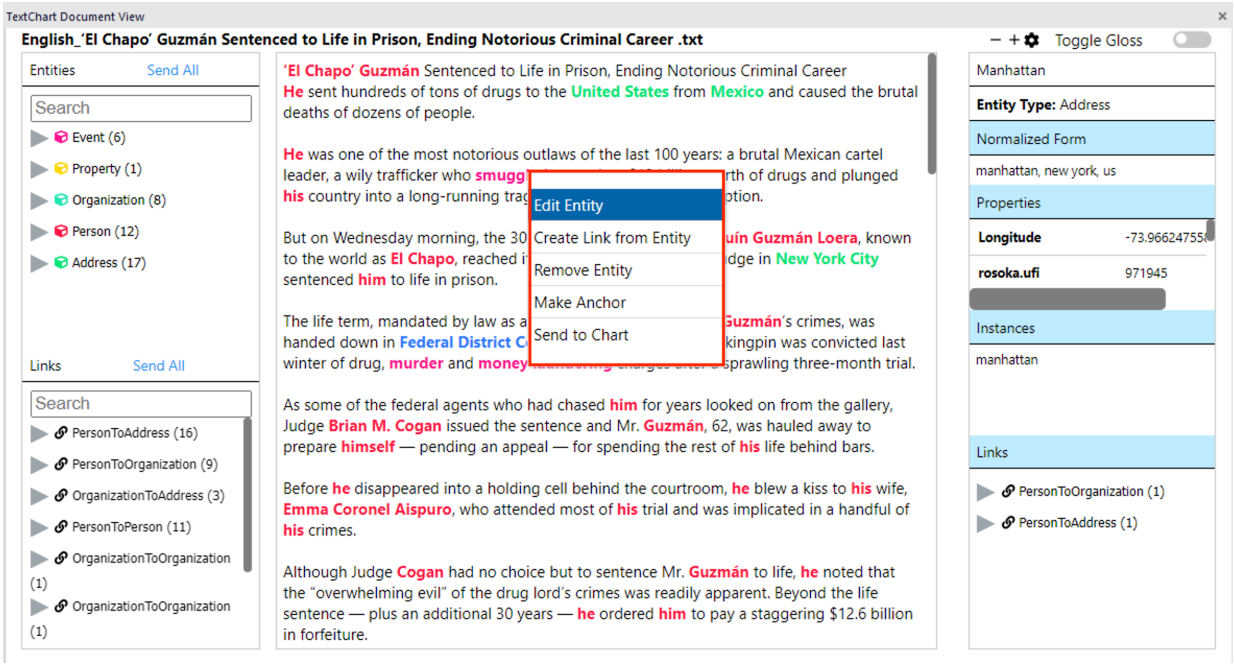
© N. Harris Computer Corporation



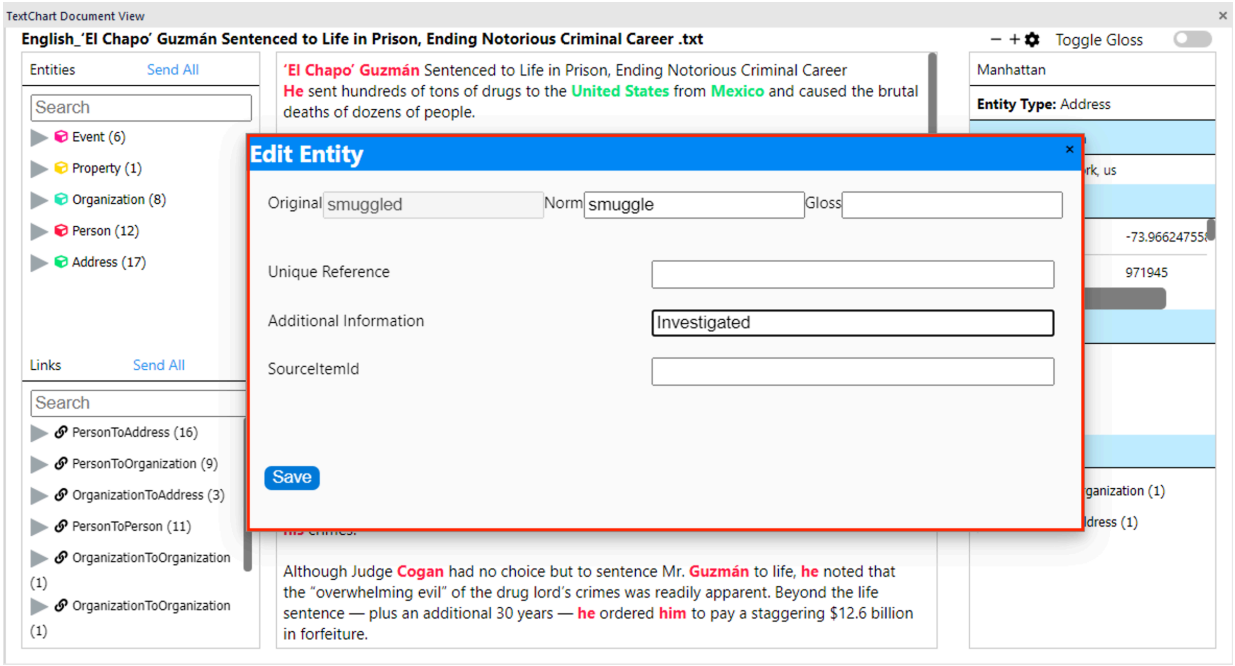
Editing results

If TextChart identified an entity or a link successfully, but the extracted result isn't exactly how you want it, you can right-click the highlighted text in the Document View and select **Edit Entity**.

© N. Harris Computer Corporation

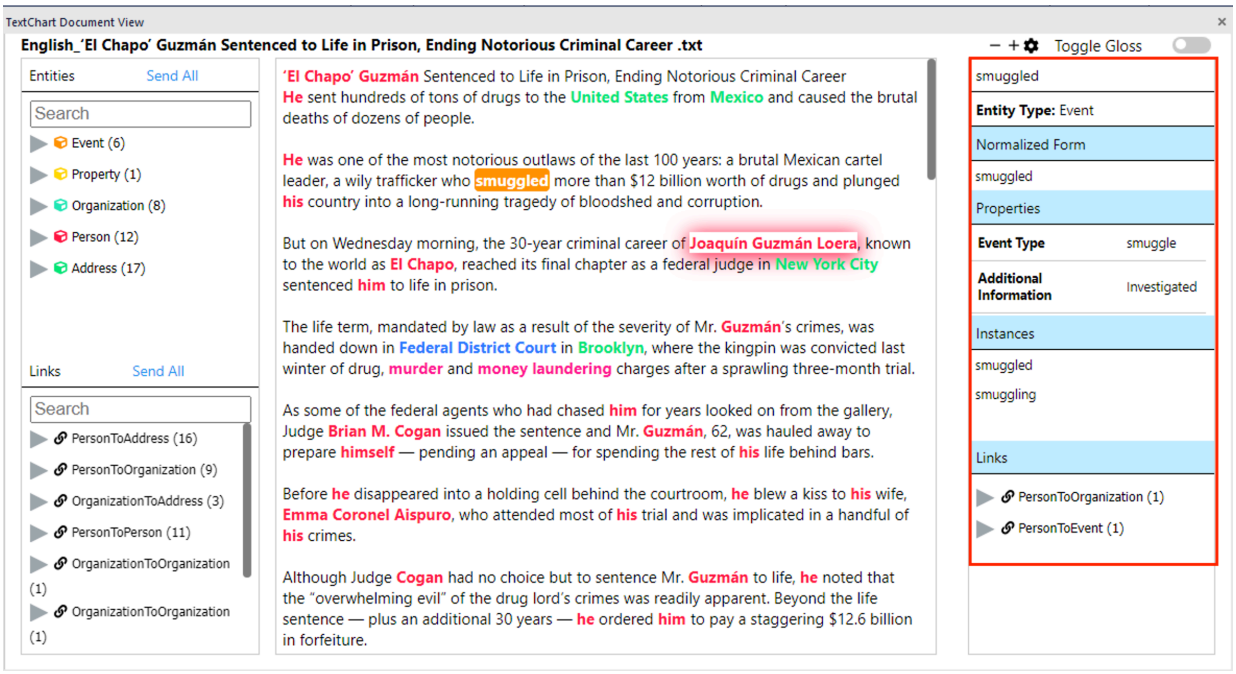


The behavior of the tool is the same when you're editing results as it is when you're adding them. You can edit the **Norm** field, add an English **Gloss**, change the **Type**, and modify attribute information.



© N. Harris Computer Corporation

When you finish editing, the changes you made are reflected in the feature bar on the right of the Document View.



Important: After you edit an extracted entity or link, you must reprocess the other documents in the collection so that they receive the same modifications.

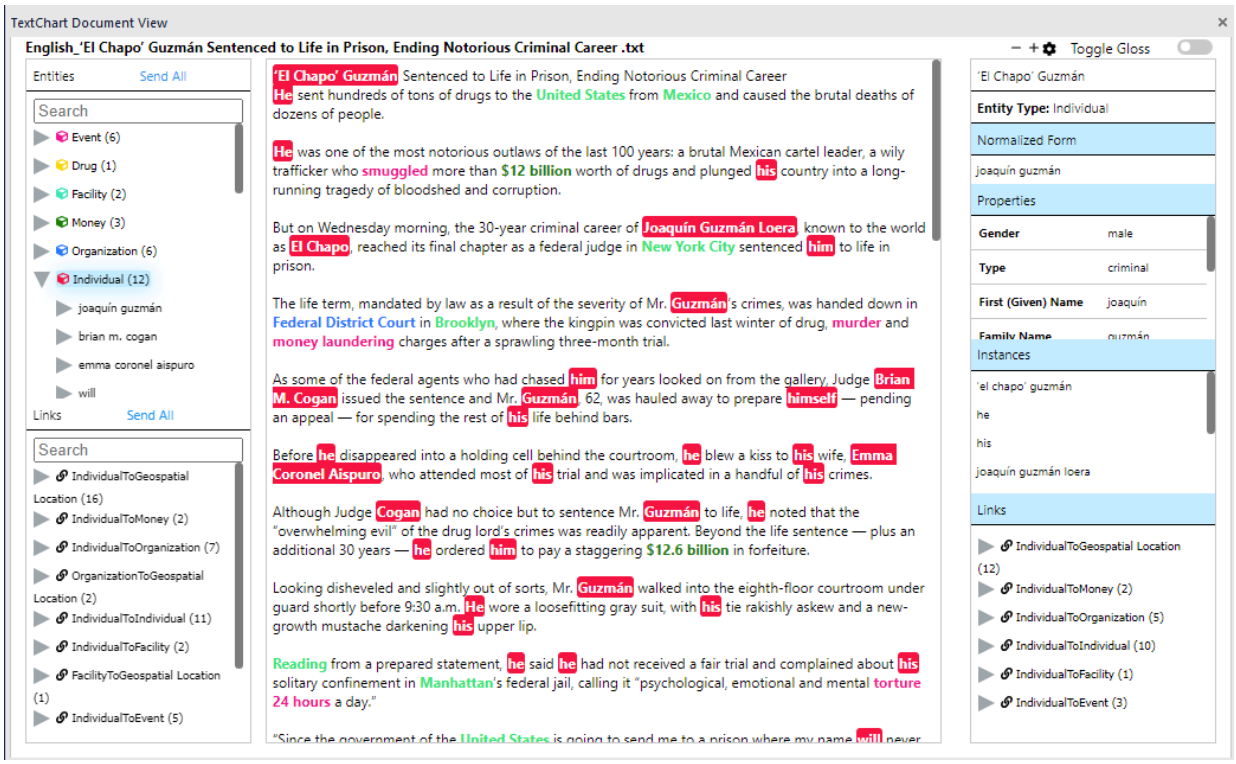
Analyzing extraction results

After i2 TextChart processes a document, it presents its results in the Document View for you to assess. The views that it presents for extracted entities and extracted links are slightly different from each other.

Viewing entities

When you select an entity type in the **Entities** list, TextChart highlights the entities of that type in the document.

© N. Harris Computer Corporation

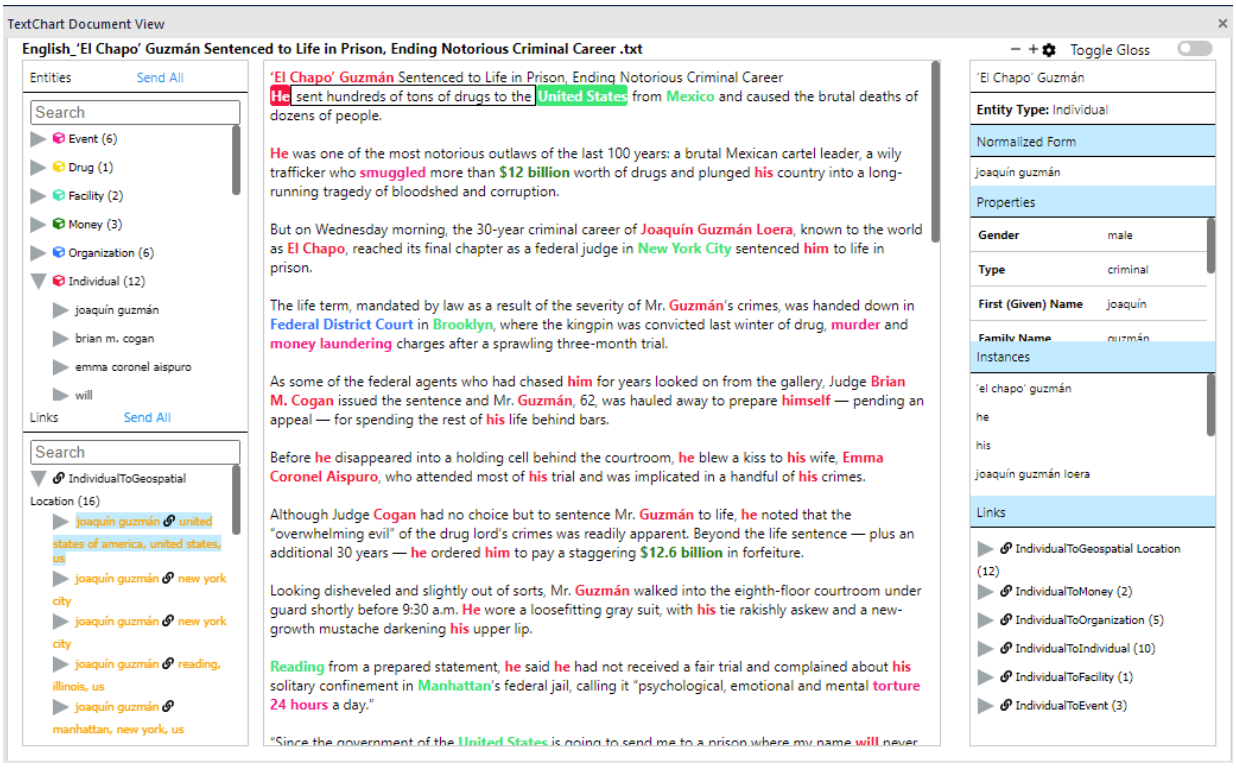


The list also displays the number of entities of each type in the document. If you expand an entity type and select a particular instance, TextChart opens the entity modification window on the right of the Document View.

Viewing links

When you select a link type in the **Links** list, TextChart highlights the entities and the predicates that contribute to links of that type.

© N. Harris Computer Corporation



As with the extracted entities, you can expand link types in the list to investigate and modify the information of any individual extracted link.

Charting extraction results

After i2 TextChart has imported and processed a collection of documents and you have curated the results, you can add the extracted entities and links to an Analyst's Notebook chart. TextChart supports sending results to the chart separately or in bulk.

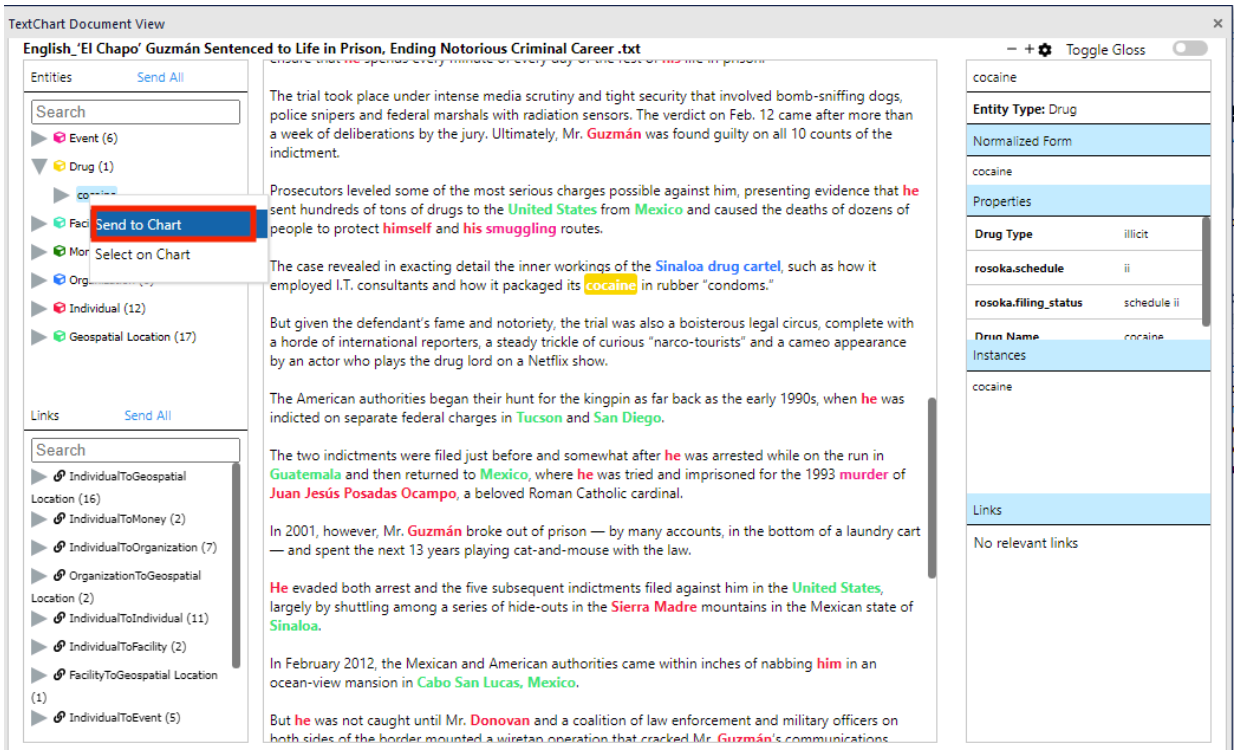
Sending entities to a chart

TextChart provides three options for sending entities to a chart. You can specify them individually, choose all entities of the same type, or simply send all the extracted entities at once.

Sending a single entity

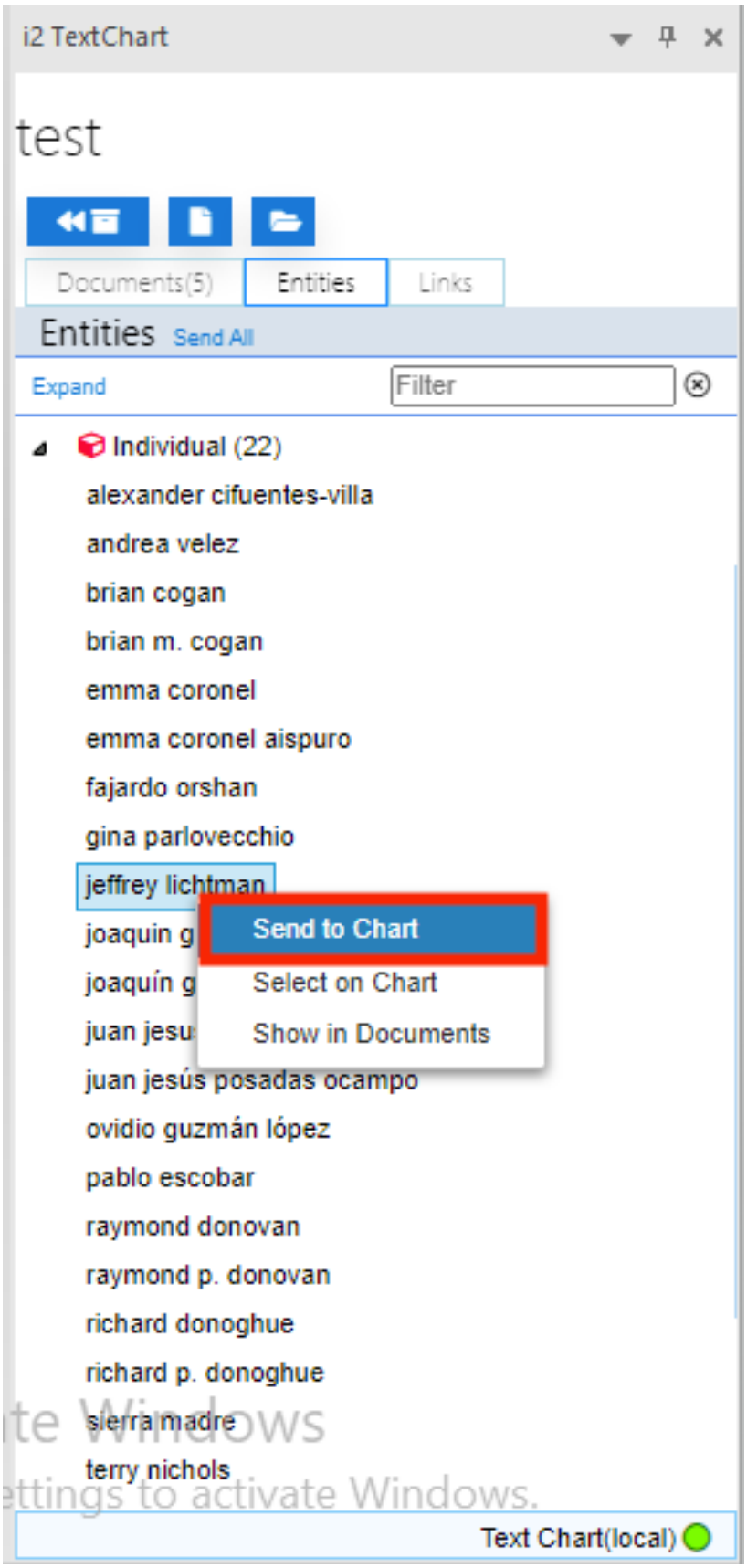
To send a single entity from the Document View to the current chart, you can use the pop-up menu for that entity in the **Entities** list on the left of the window, or the **Send To Chart** button from the entity details pane on the right of the window.

© N. Harris Computer Corporation



If an extracted entity is already on the chart, you can use the same pop-up menu to select it there.

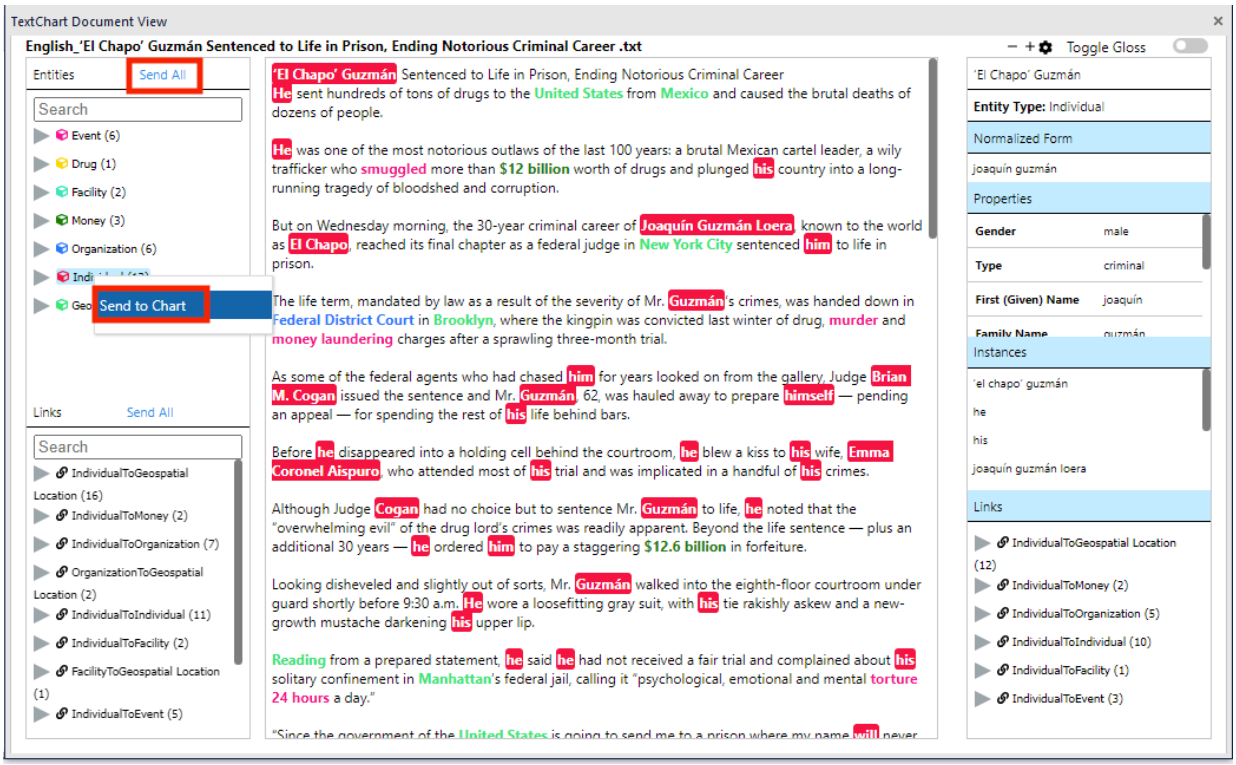
© N. Harris Computer Corporation



© N. Harris Computer Corporation

Sending several entities

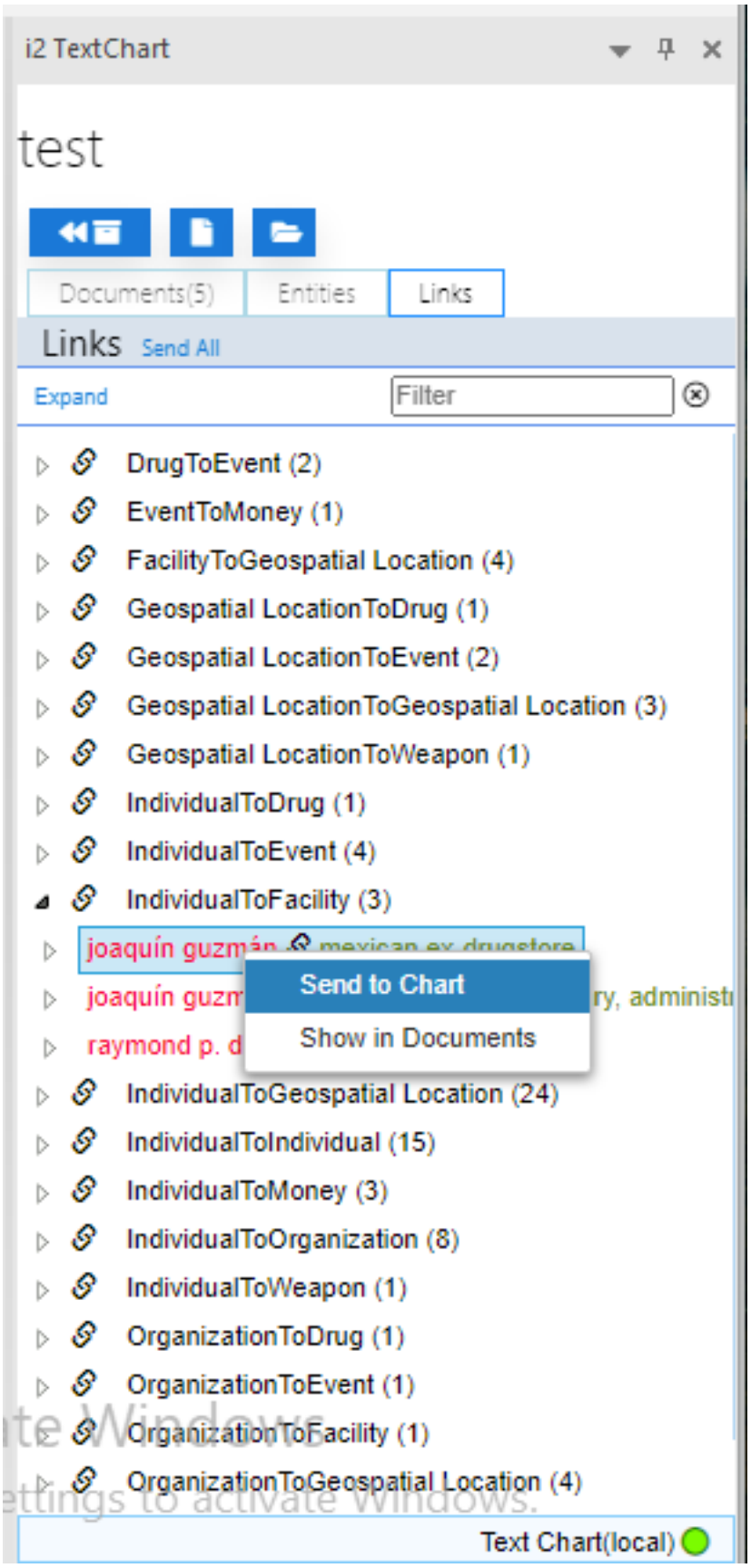
To send all the entities of the same type from the Document View to a chart, you can use the pop-up menu for that entity type in the **Entities** list. Alternatively, to send *all* the entities extracted from the document to the chart, you can click the **Send All** button at the top of the list.



Sending links to a chart

TextChart provides the same three options for sending extracted links to the chart as it does for sending entities, and you use them in the same way.

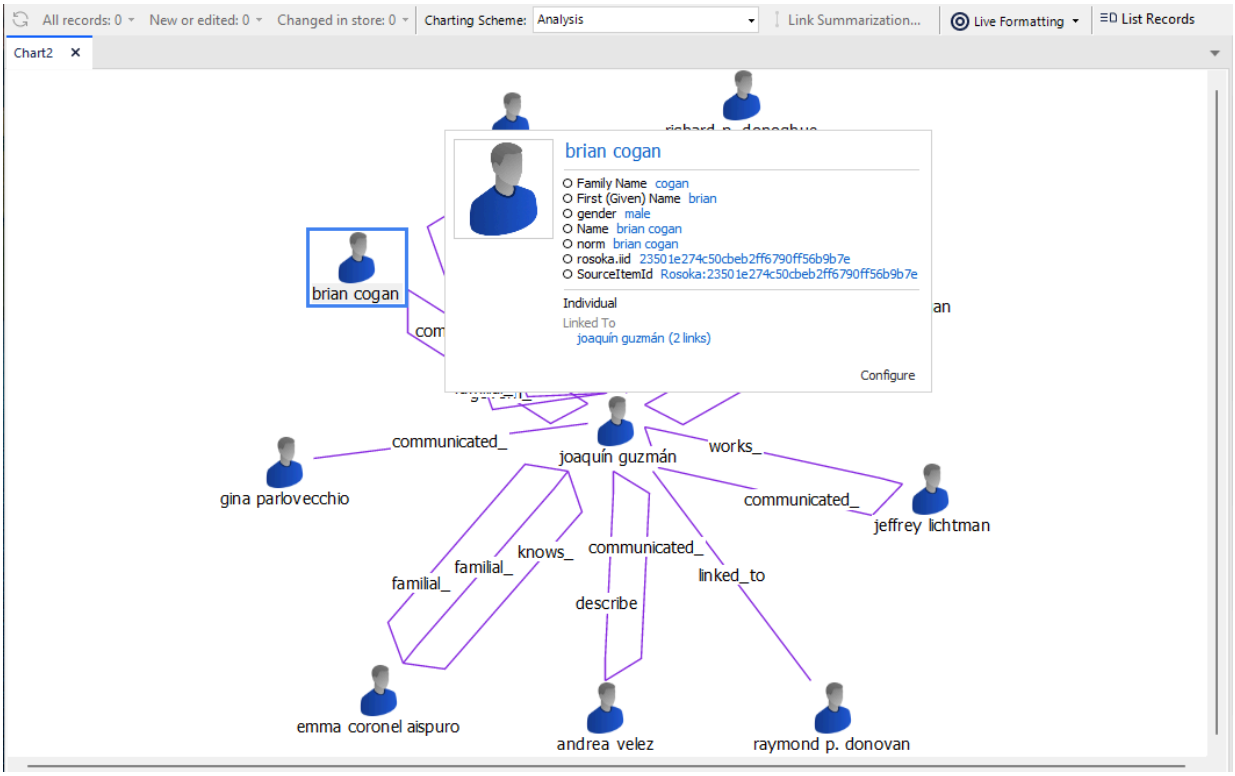
© N. Harris Computer Corporation



© N. Harris Computer Corporation

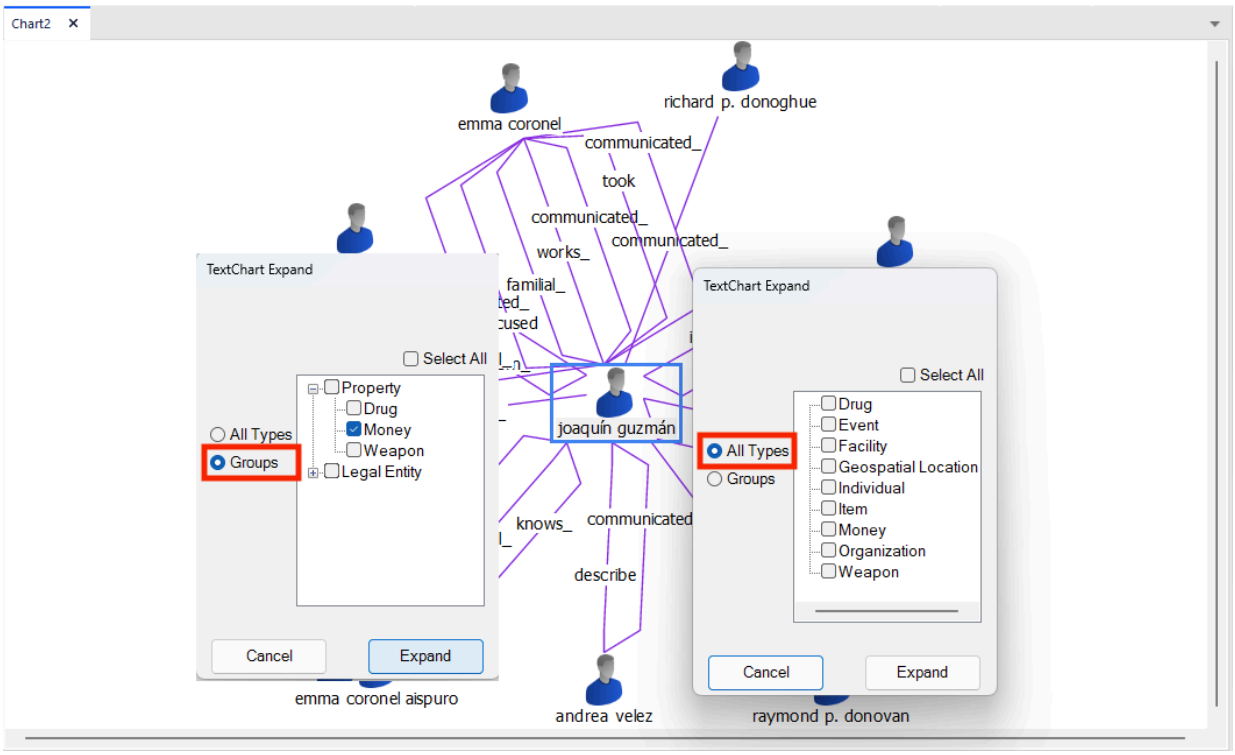
Viewing entities on the chart surface

You can hover over *any* entity on the Analyst's Notebook chart surface to display a window containing information about all the properties that are associated with that entity.



After you've added an entity *from TextChart* to the chart, you can use it to start adding related entities from the other documents that you've processed. Right-click the entity, and select **Rosoka Expand**.

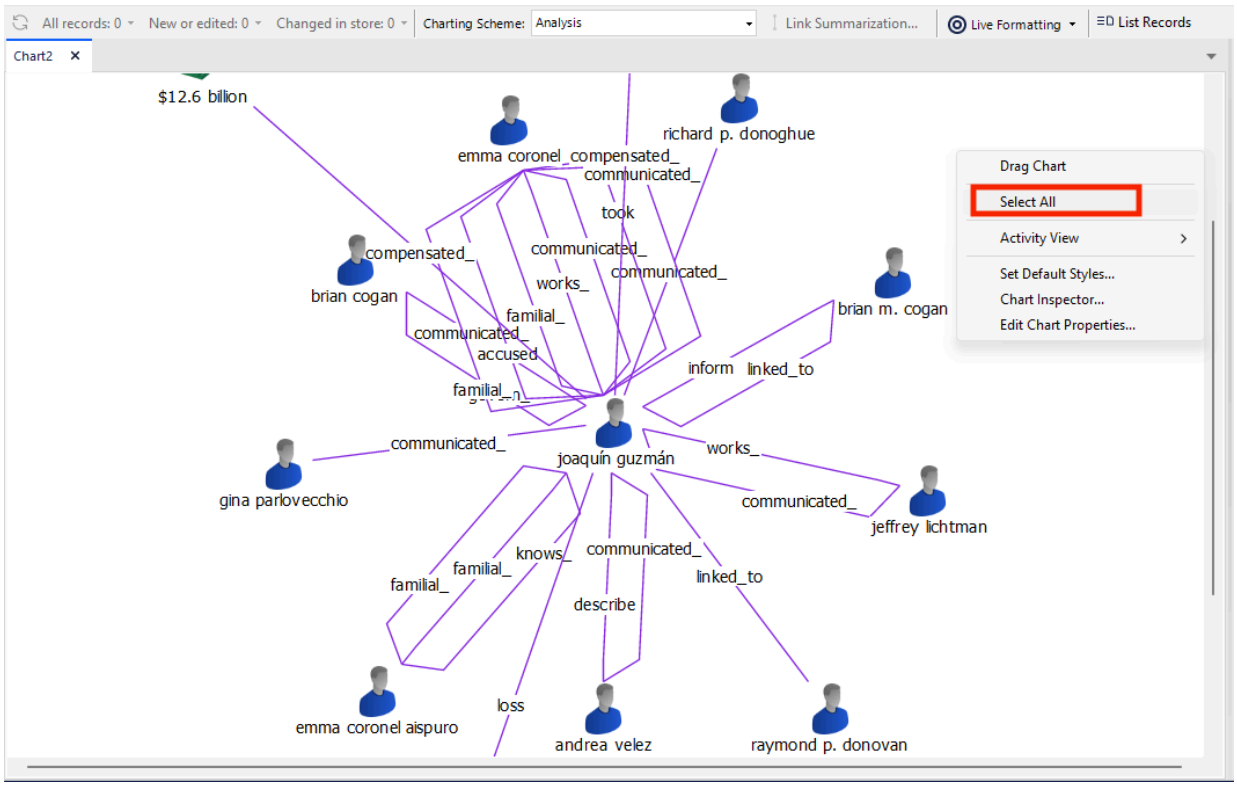
© N. Harris Computer Corporation



You can choose to add related entities of particular types, or groups of types; or you can use **Select All** to add all related entities in a single action.

When a chart contains multiple TextChart entities, you can perform the same **Rosoka Expand** operation on all of them by right-clicking the chart surface and choosing **Select All** first.

© N. Harris Computer Corporation

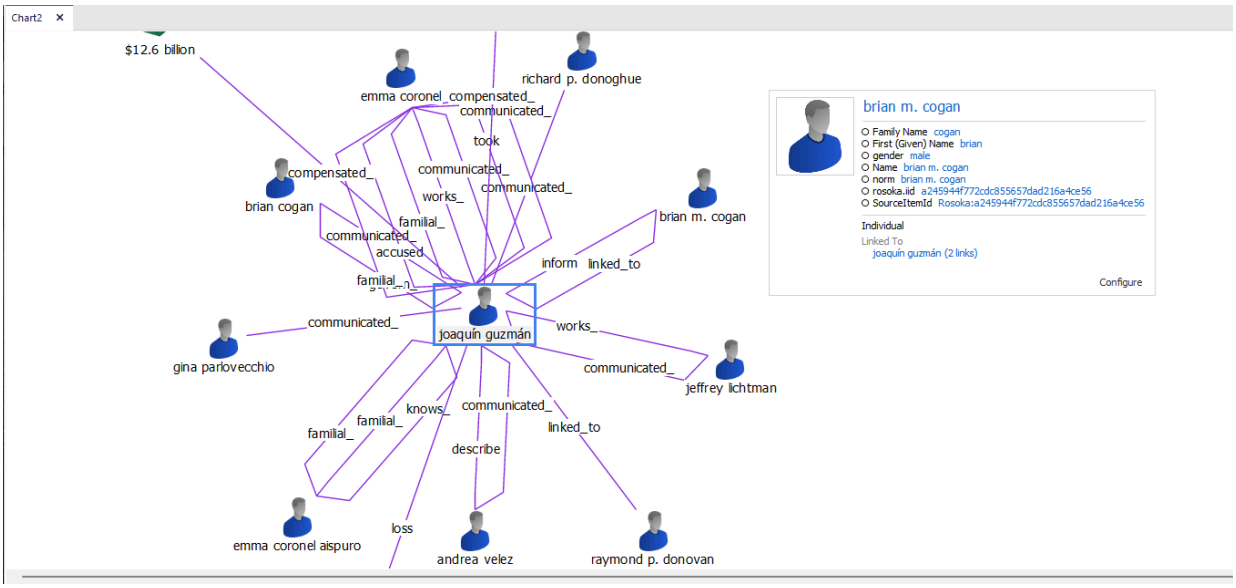


Viewing links on the chart surface

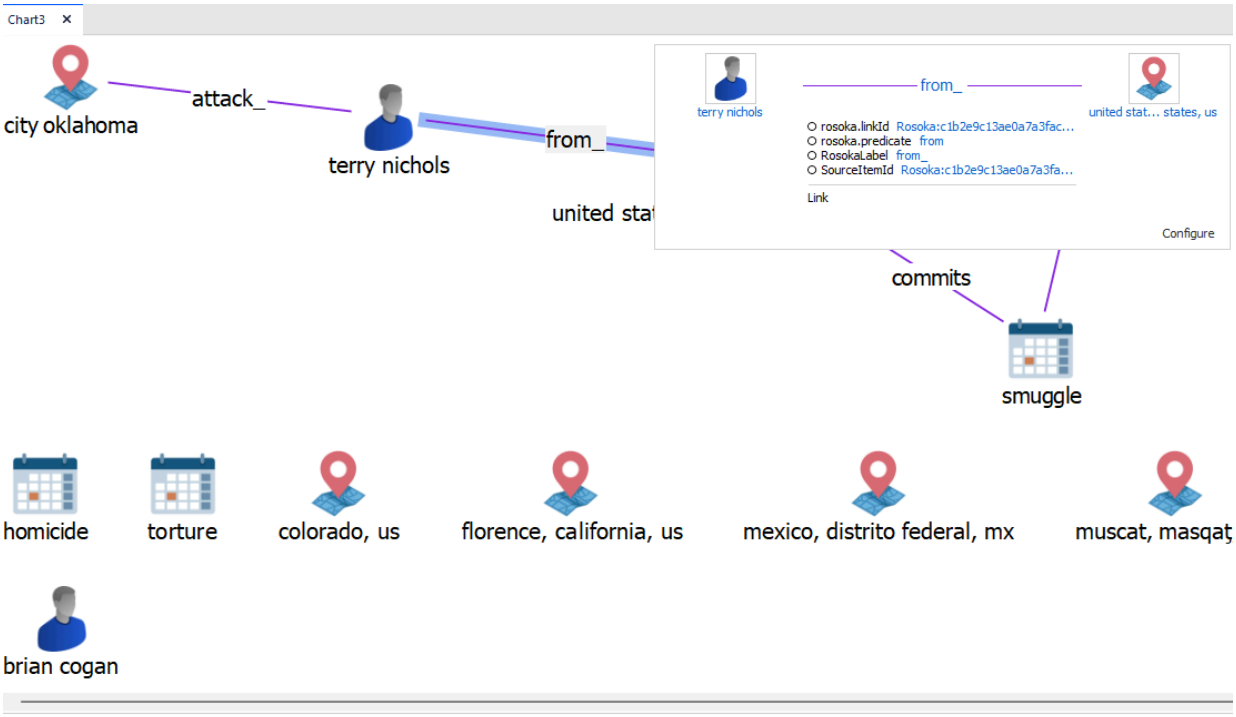
After you've added all the entities to the chart that you want to add, you can start to investigate the relationships between them, including the predicates that were responsible for the links being identified in the first place.

For example, hovering over a connected entity reveals a list of the entities that it is linked to, in addition to its properties.

© N. Harris Computer Corporation



Similarly, you can hover over a TextChart link to highlight the connection between two entities and display additional extraction information.

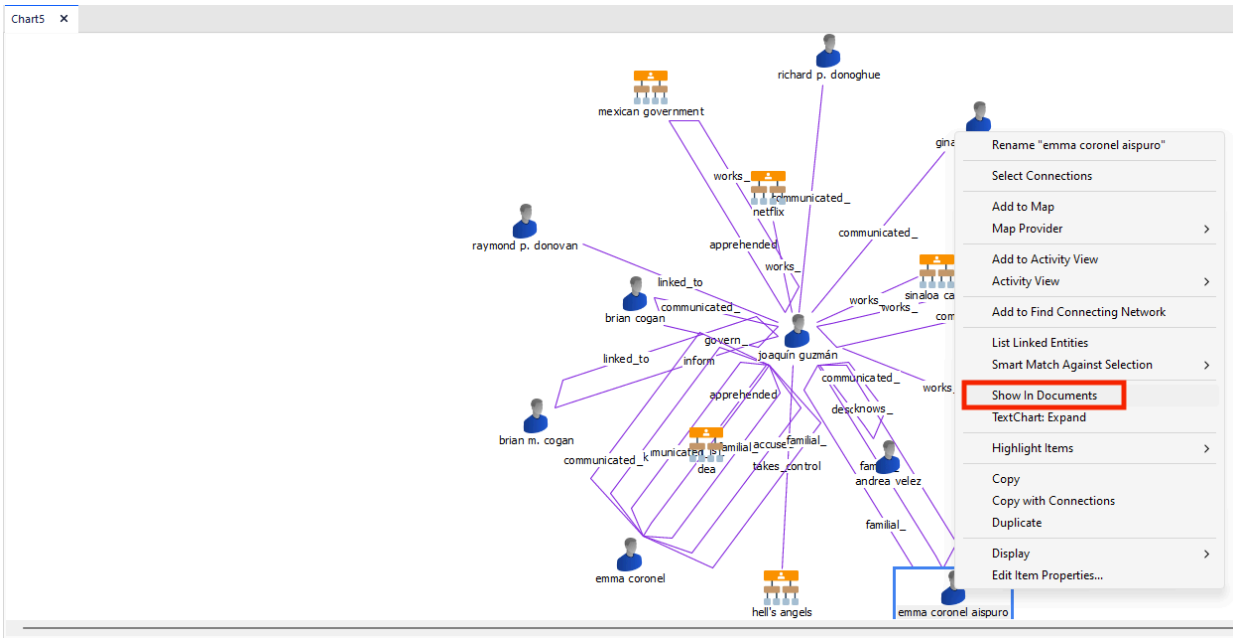


© N. Harris Computer Corporation

Listing source documents

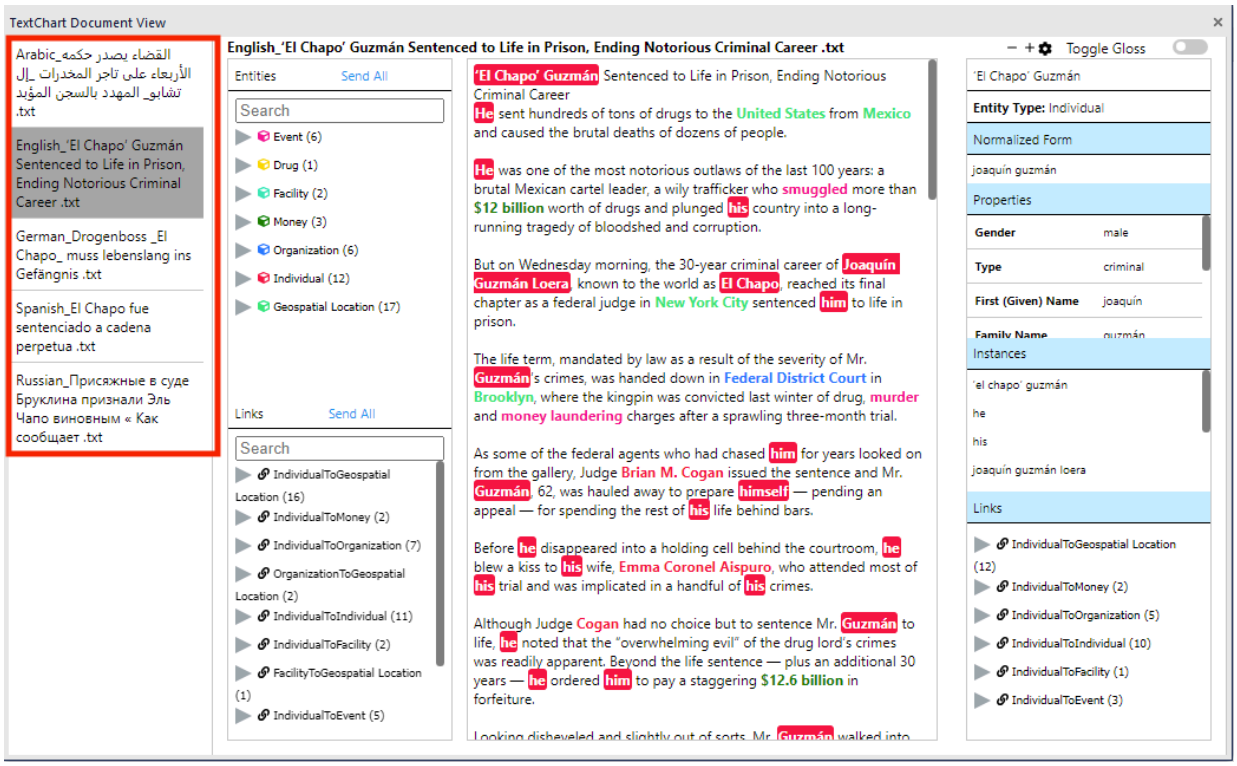
For any entity or link that was added to the Analyst's Notebook chart surface through TextChart, you can quickly refer back to the documents from which it was extracted.

To retrieve all the documents for an entity or a link, right-click and select **Show In Documents** to reopen the Document View.



On the left of the view, the list of documents allows you to select which document appears on the right. Each document appears with text highlighting that reveals the entity or link in question.

© N. Harris Computer Corporation



Mapping to i2 Analyze types

When Analyst's Notebook is connected to an i2 Analyze server, the types of the entities and links that users can add to a chart (and the types of the properties that they can contain) are governed by a *schema*. The schema has an associated *charting scheme* that controls the appearance of entities and links.

Individual deployments of i2 Analyze have custom schemas that define entity and link types that are appropriate to the analysis they perform. Different organizations will use different types: even if two organizations used entity types named Person, those types would have different definitions.

TextChart, on the other hand, has a fixed list of types that it can extract from documents, which is defined in the LxBase. In order for Analyst's Notebook to receive entities and links from TextChart, and to display and analyze them successfully, there must be a mapping from the TextChart types to an i2 Analyze schema.

Default type mapping

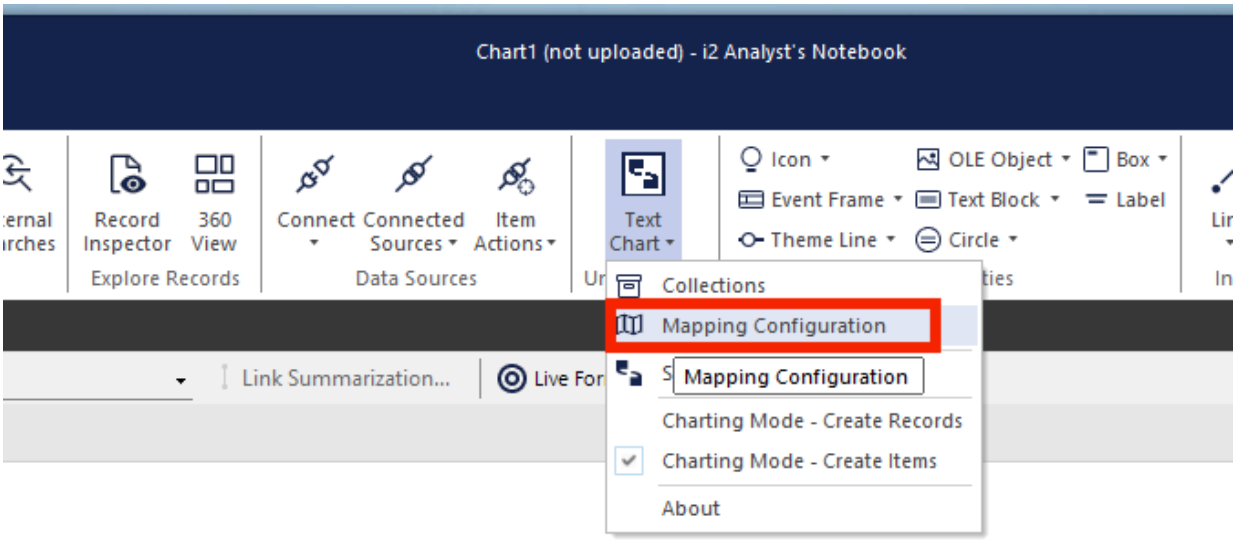
When you install i2 TextChart, it includes an i2 Analyze schema that defines all the same types that the LxBase defines, and a *mapping configuration* that provides a one-to-one mapping between the two sets of definitions. By default, TextChart uses that schema and mapping configuration when you send extracted entities and links to the chart.

Through Analyst's Notebook, you can make changes to the default mapping configuration so that some types are extracted differently - or so that they're not extracted at all. You can also create

© N. Harris Computer Corporation

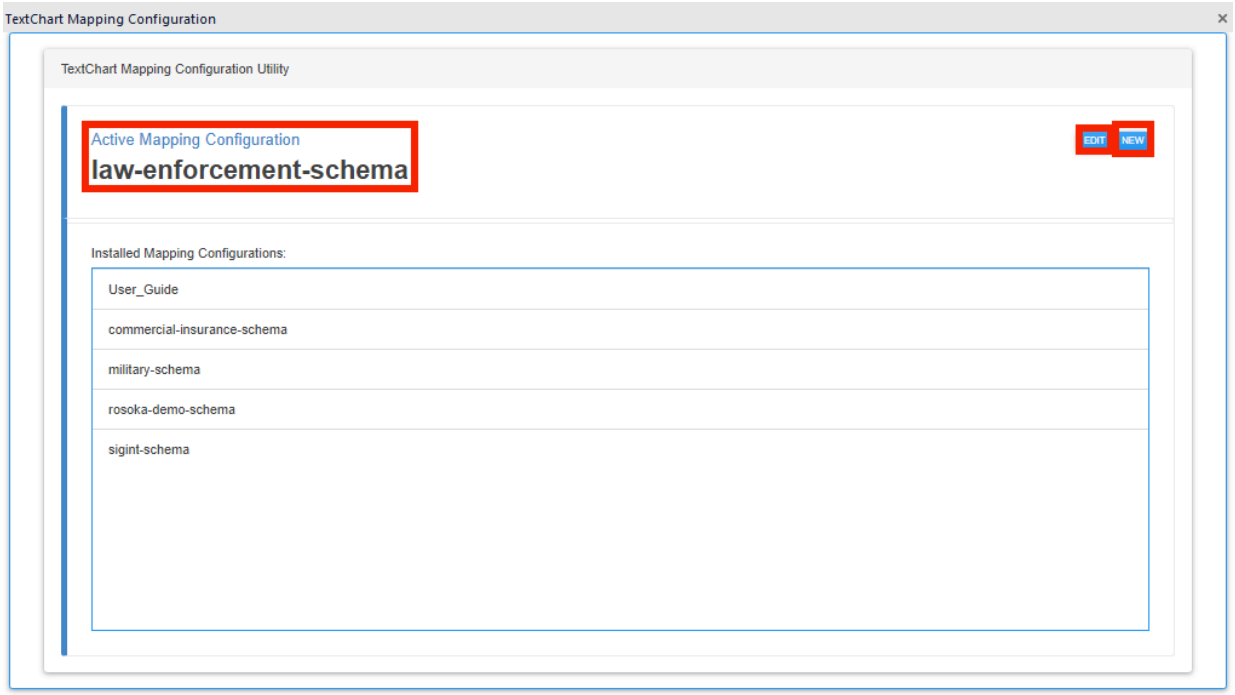
mapping configurations for different i2 Analyze schemas, so that the entities and links you extract from documents are aligned with those that you're already using.

To start the process, select **Text Analytics > Mapping Configuration**.



The command opens the Mapping Configuration window, where you can make changes to the active configuration, switch to a different configuration, or create a mapping configuration of your own.

© N. Harris Computer Corporation

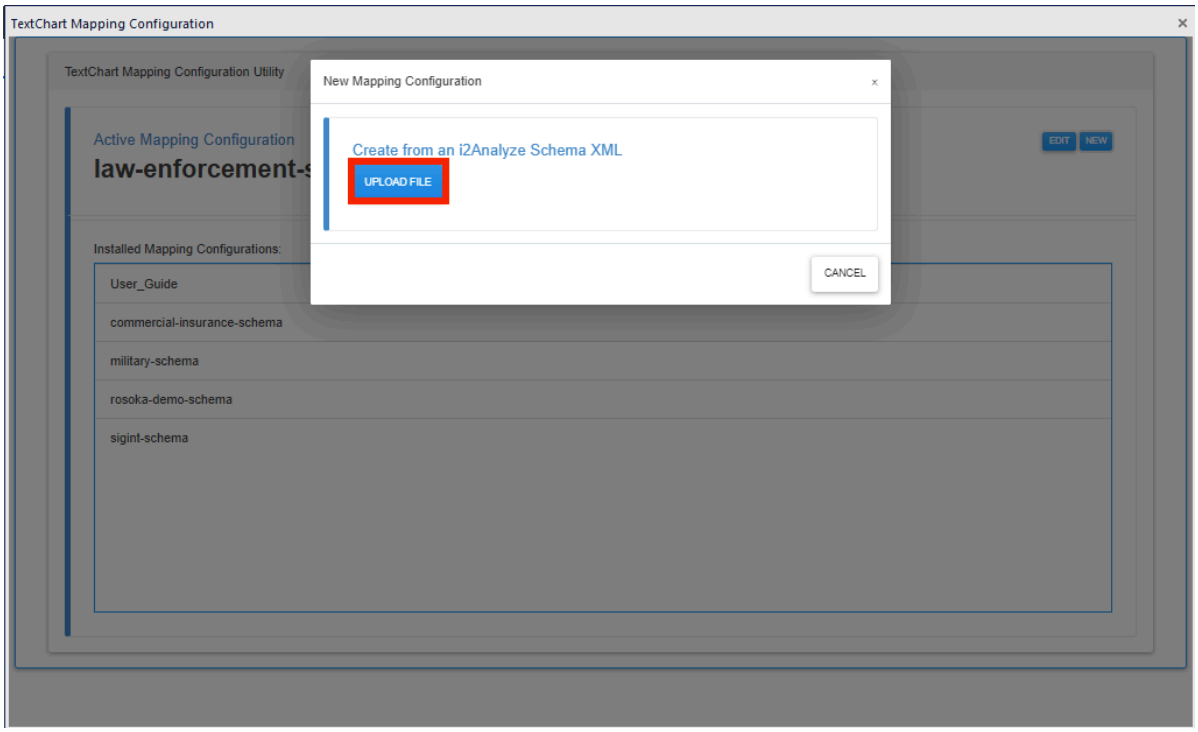


Mapping to a different schema

To start the process of mapping TextChart types to a different i2 Analyze schema, you must tell TextChart about the schema that you want to map to!

1. Click **New** to open the New Mapping Configuration window.

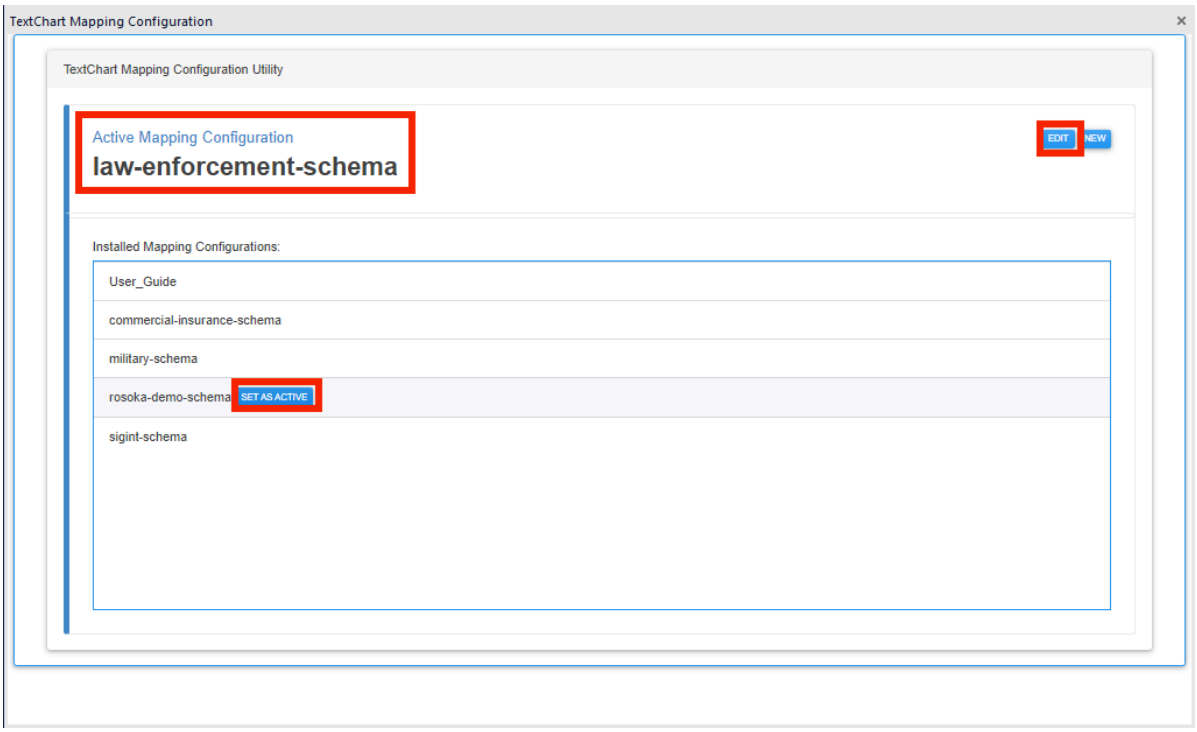
© N. Harris Computer Corporation



2. Click **UPLOAD FILE**, and locate the XML file that contains the i2 Analyze schema you want to use.

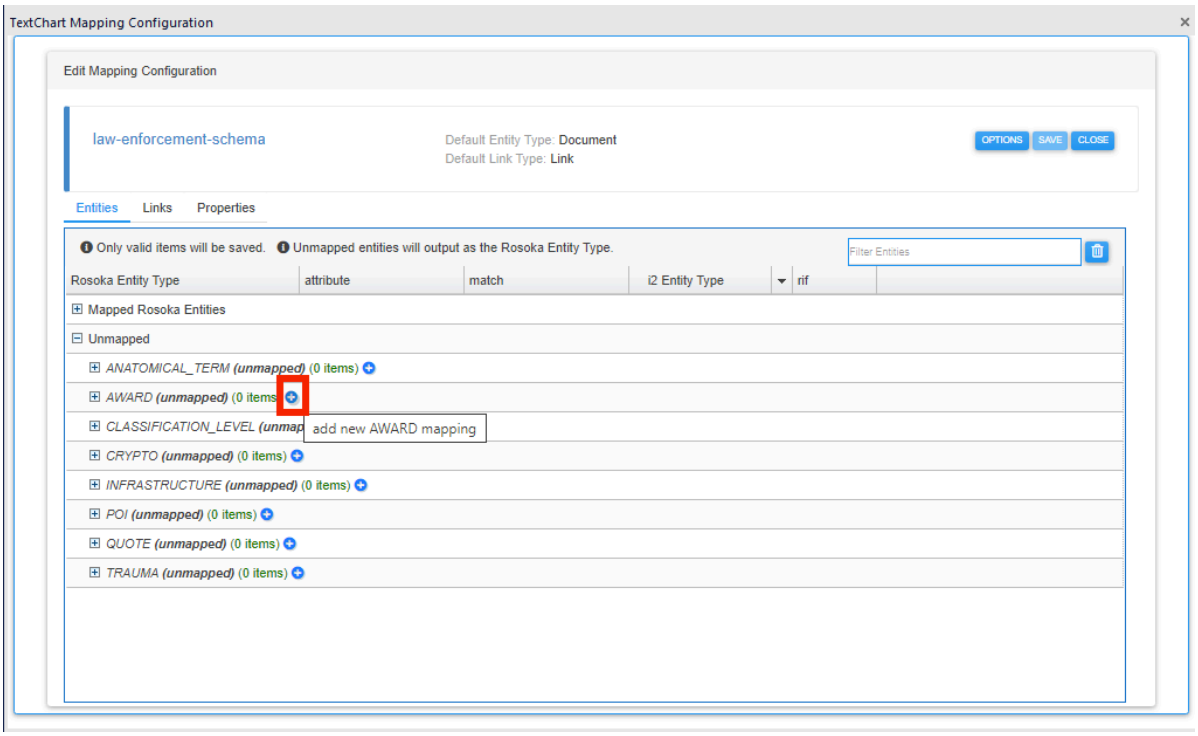
   When the upload succeeds, click **OK** to close the dialog. TextChart creates an empty mapping configuration for the schema and makes it active.

© N. Harris Computer Corporation



3. Click **Edit** to open the new mapping configuration. At this stage, all of the TextChart types are "unmapped".

© N. Harris Computer Corporation



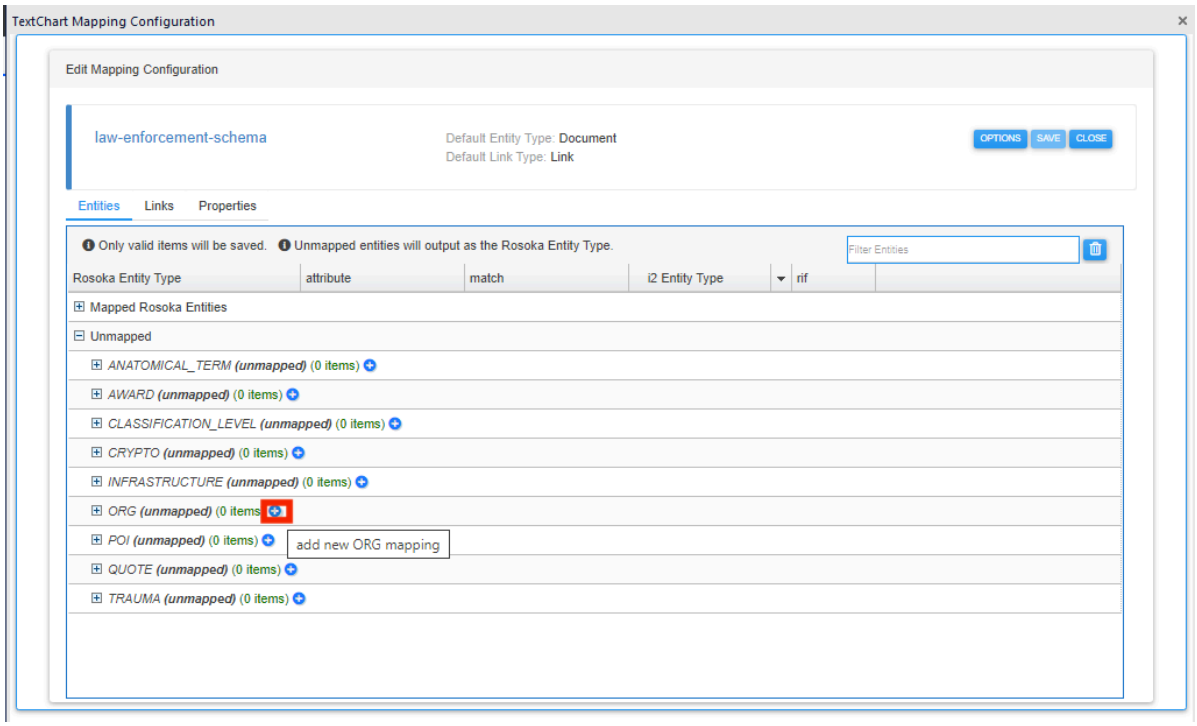
Note: When a configuration contains no mapping for a particular TextChart entity or link type, TextChart falls back to using the default mapping (and the default schema) for that type.

Mapping entity types

At this point, you can start to create the mappings from TextChart types to i2 Analyze schema types. It's usually easier to start with the entity types.

1. Click the "plus" icon next to the name of the TextChart type that you want to map. The type moves from the "unmapped" list to the "mapped" list:

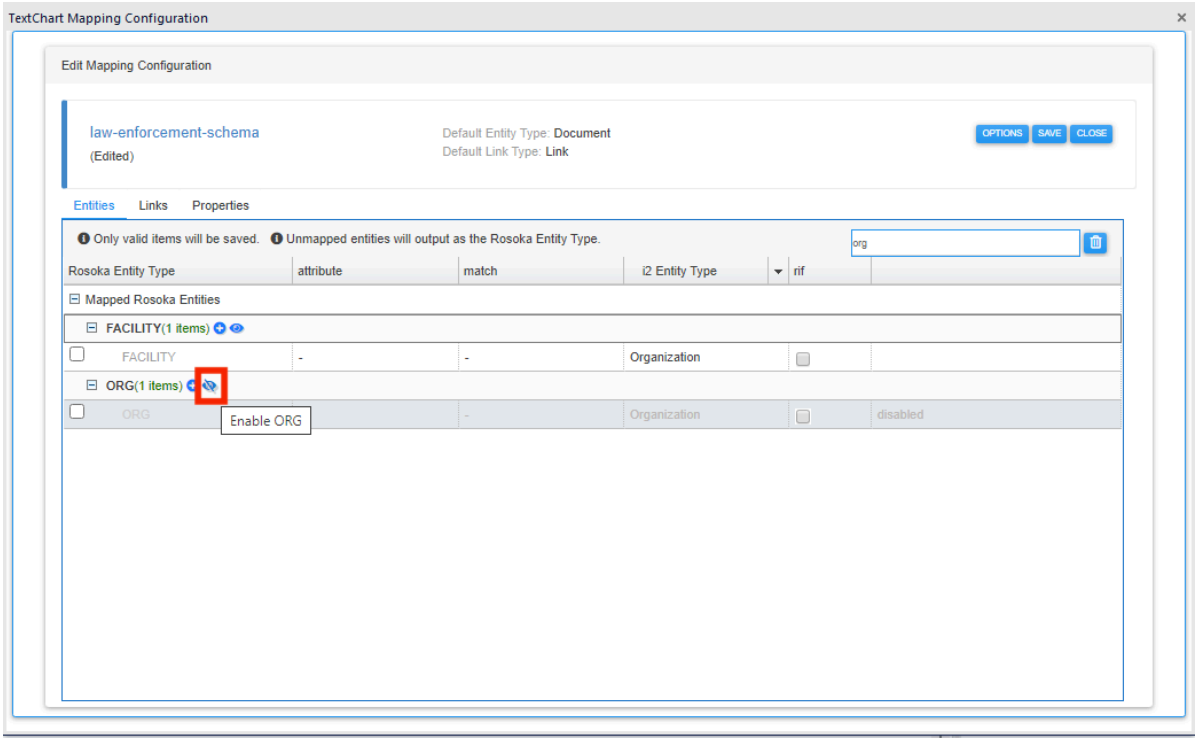
© N. Harris Computer Corporation



When you first move a type to the "mapped" list like this, the mapping - and, effectively, extraction - is disabled. Entities with this type do not appear in TextChart results.

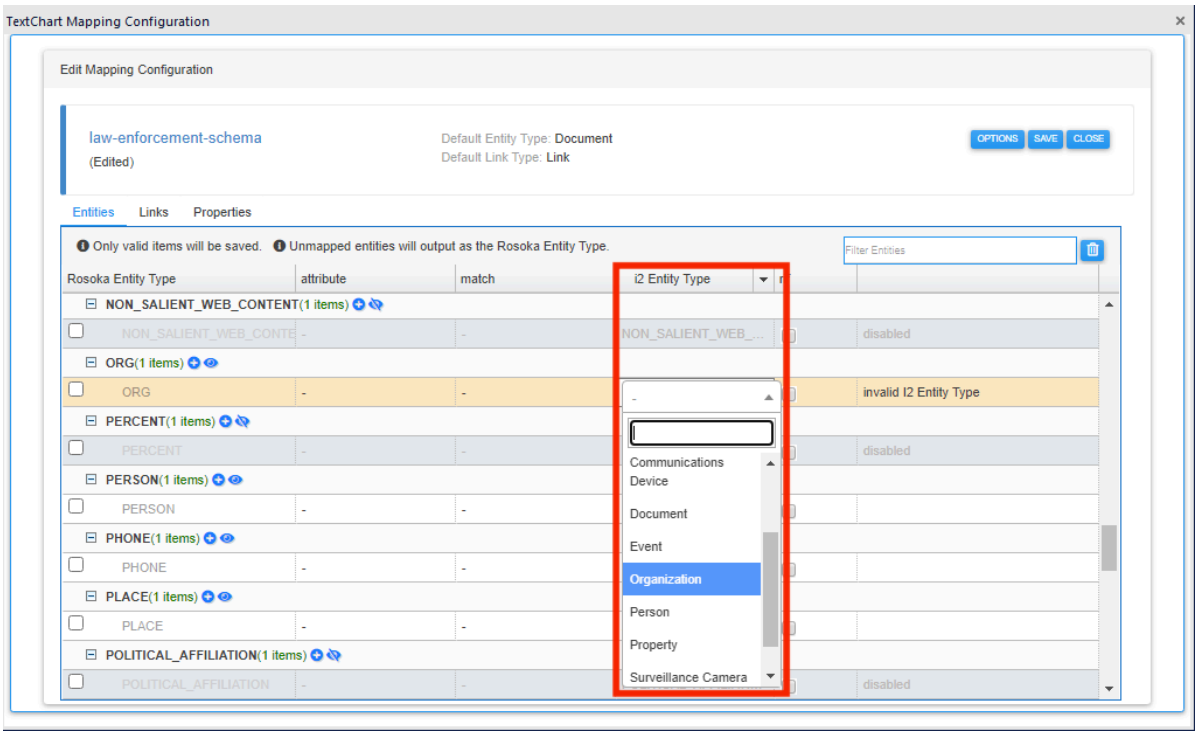
2. To enable the mapping, click the crossed-out "eyeball" icon next to the type name.

© N. Harris Computer Corporation



3. Now, you can choose the entity type in the i2 Analyze schema that you want the TextChart type to map to. Open the drop-down list in the **i2 Entity Type** column.

© N. Harris Computer Corporation



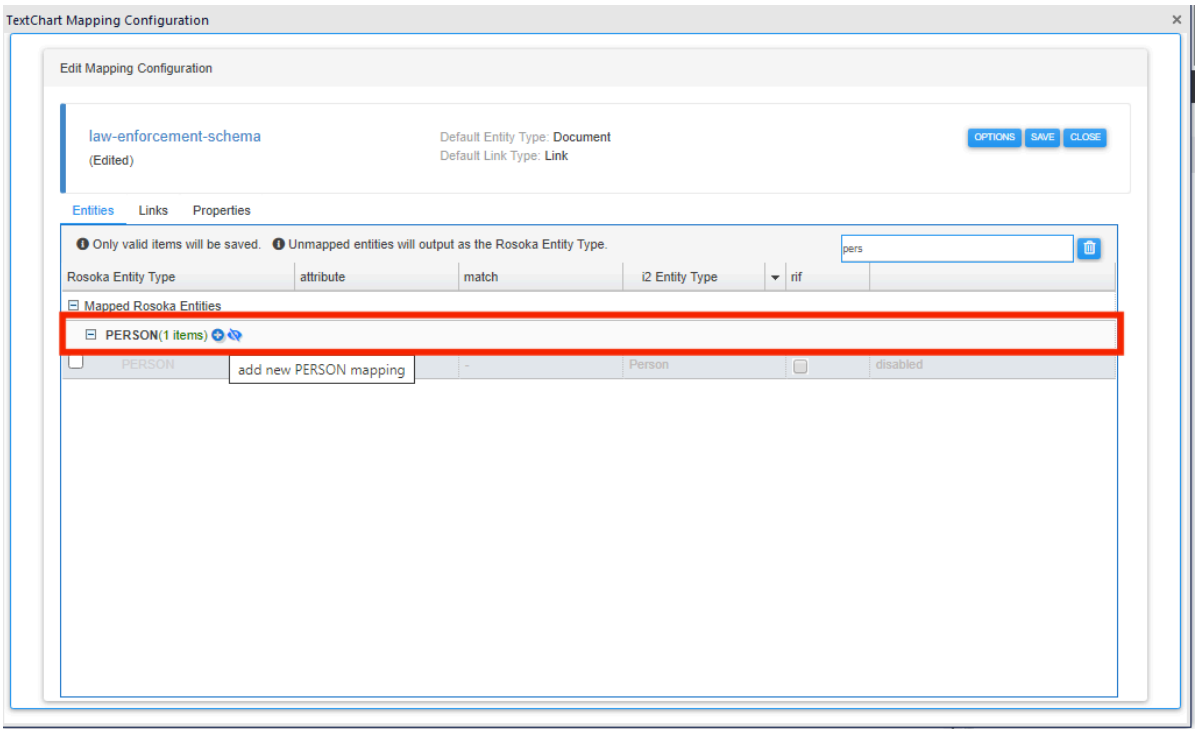
4. Select the i2 entity type that you want to use, and then click **Save** to update the mapping configuration.

Mapping entity types based on attribute values

Sometimes, you might want to map a TextChart type to different i2 Analyze types depending on the value of a particular attribute. For example, if the i2 Analyze schema has separate Male and Female entity types, you might map TextChart Person entities whose Gender attribute is "Male" to the former, and those whose Gender attribute is "Female" to the latter.

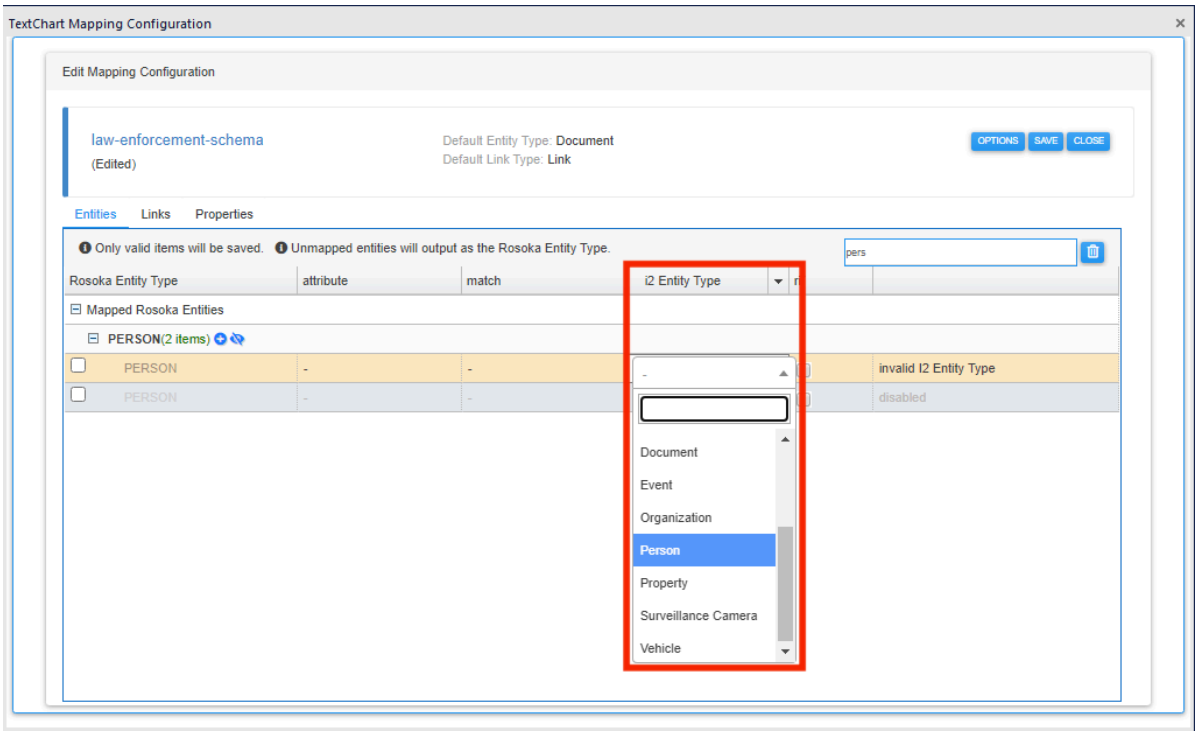
1. As before, click the "plus" icon next to the name of the TextChart entity type that you want to move to the "mapped" list.
2. Click the "plus" icon again to add a new row to the mapping configuration for the same TextChart entity type.

© N. Harris Computer Corporation



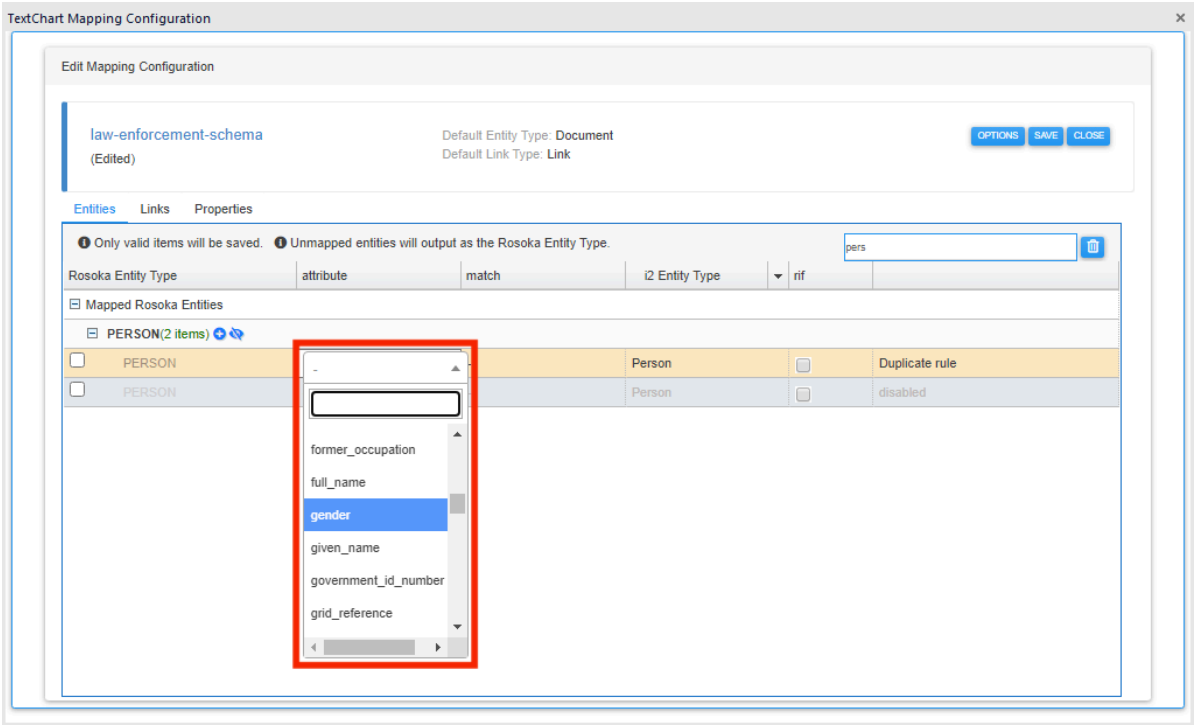
3. In the **i2 Entity Type** column, select one of the types that the TextChart entity type will be mapped to.

© N. Harris Computer Corporation



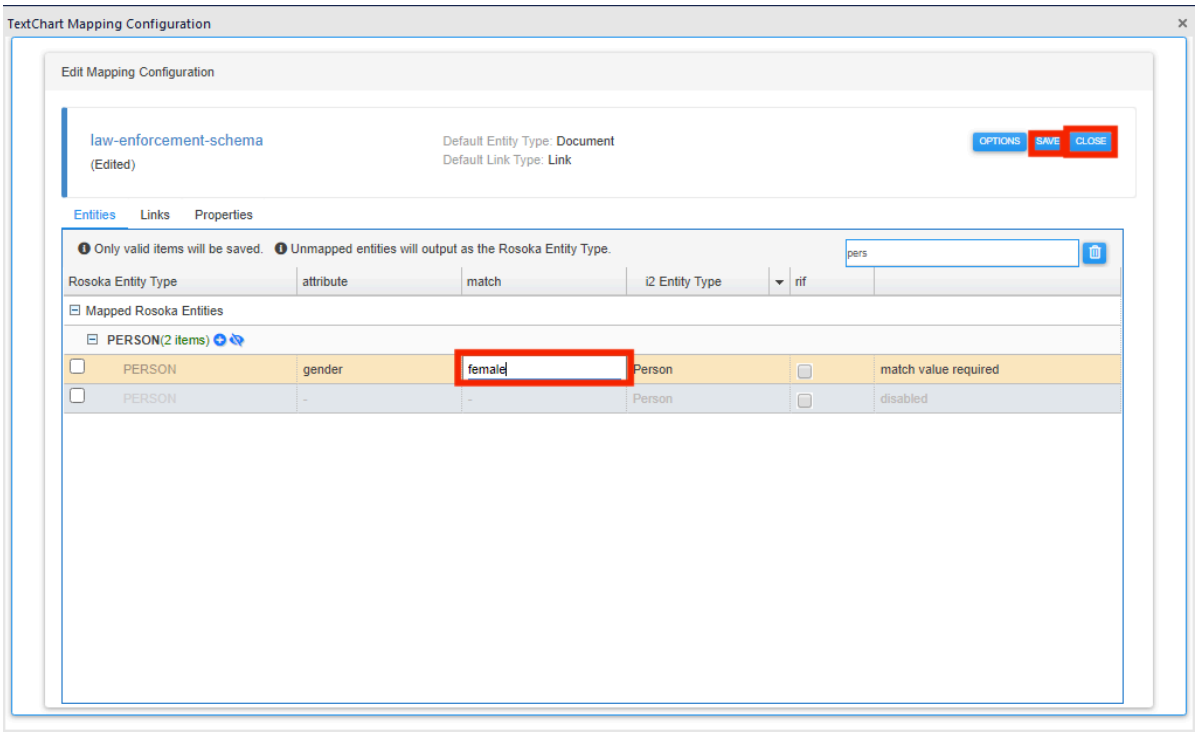
4. In the **attribute** column, select the TextChart attribute whose value will control whether this mapping is used.

© N. Harris Computer Corporation



5. In the **match** column, specify the attribute value that will cause this mapping to be used.

© N. Harris Computer Corporation



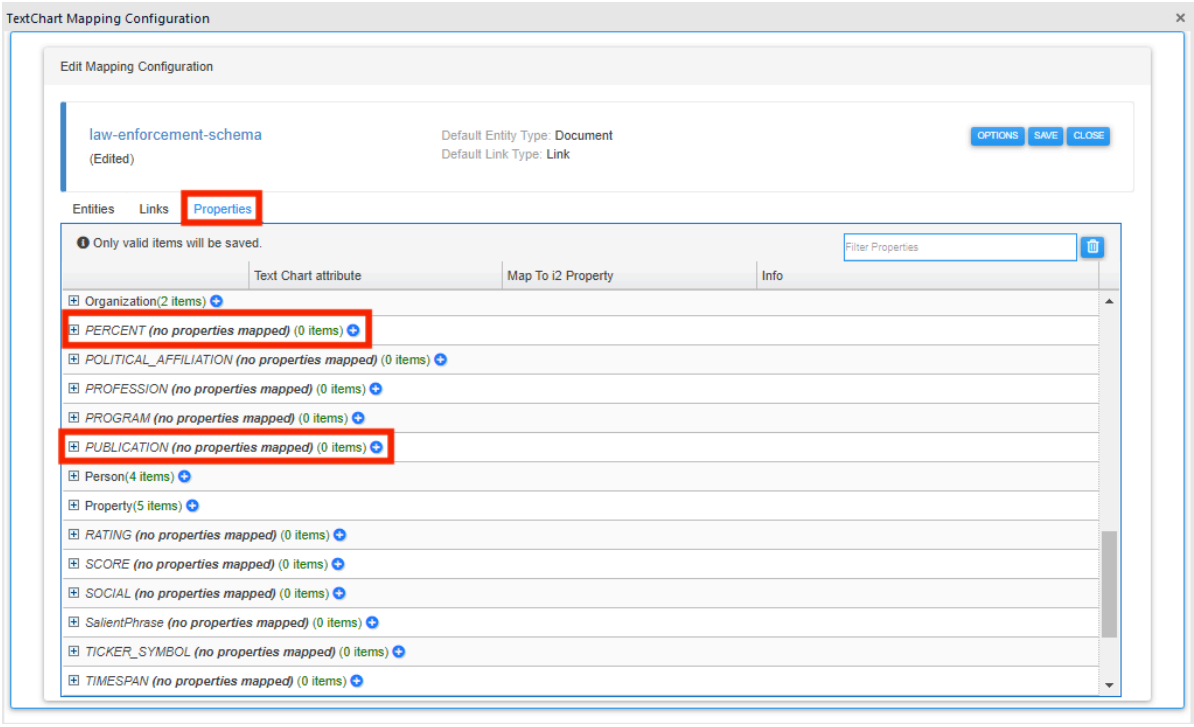
When you save this mapping configuration, the entities that you send to charts from TextChart extraction results will have different types, according to the criteria that you defined.

Mapping TextChart attributes to i2 Analyze property types

After you set up mappings between TextChart and i2 Analyze entity types, the next step is to configure the mappings between their attributes and property types.

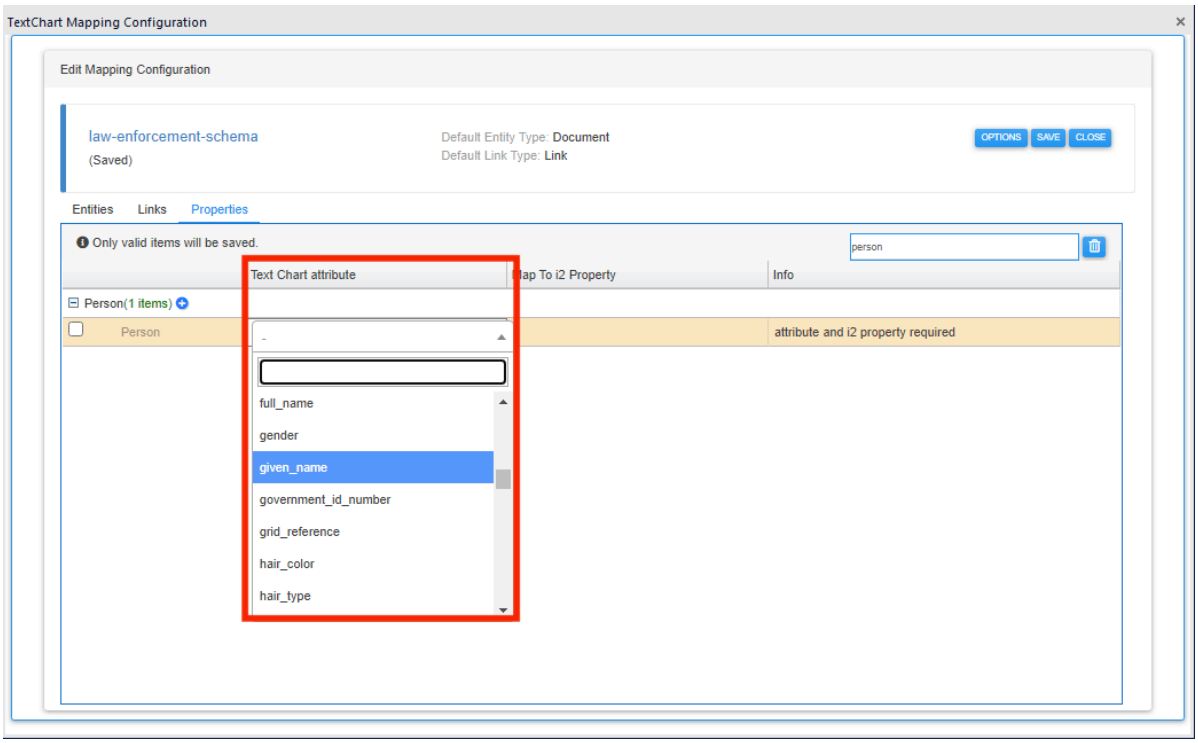
1. Select the **Properties** tab in the Mapping Configuration window to display a list of all the entity types that you've mapped.

© N. Harris Computer Corporation



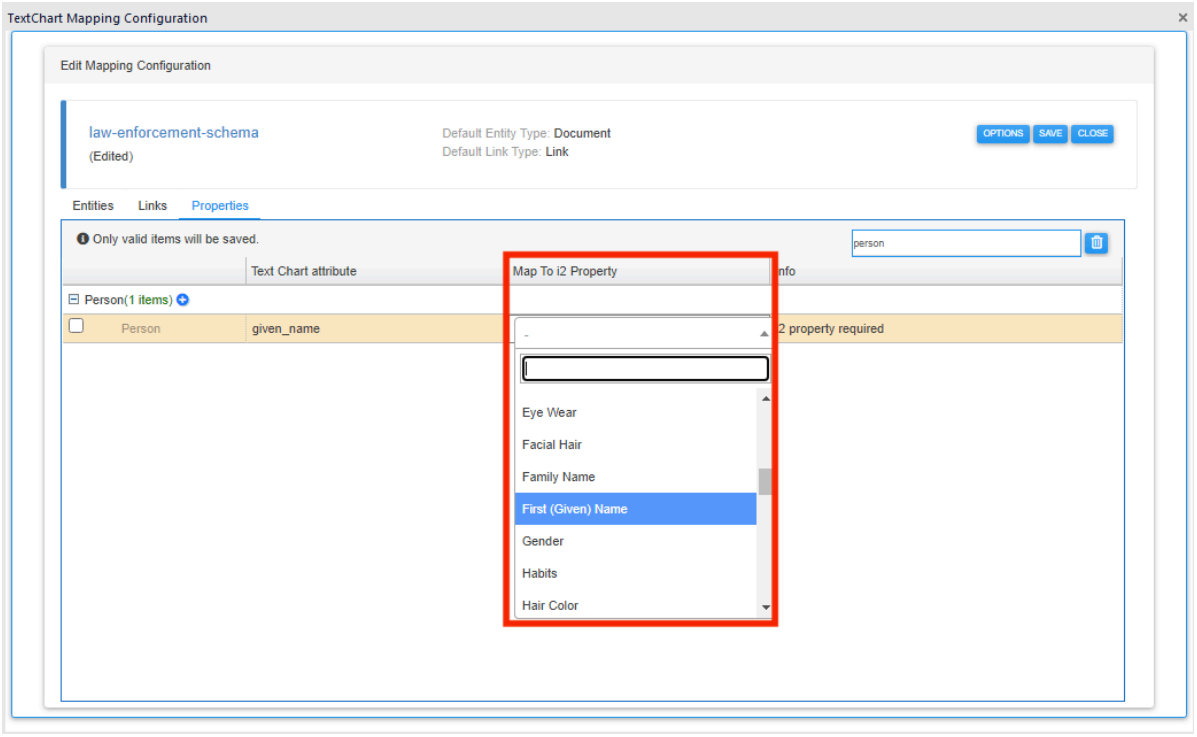
2. Click the "plus" icon next to the entity type to which you want to add a property type mapping. A new row appears below the entity type.
3. In the **TextChart attribute** column, select the attribute of the TextChart type for which you want to define a mapping.

© N. Harris Computer Corporation



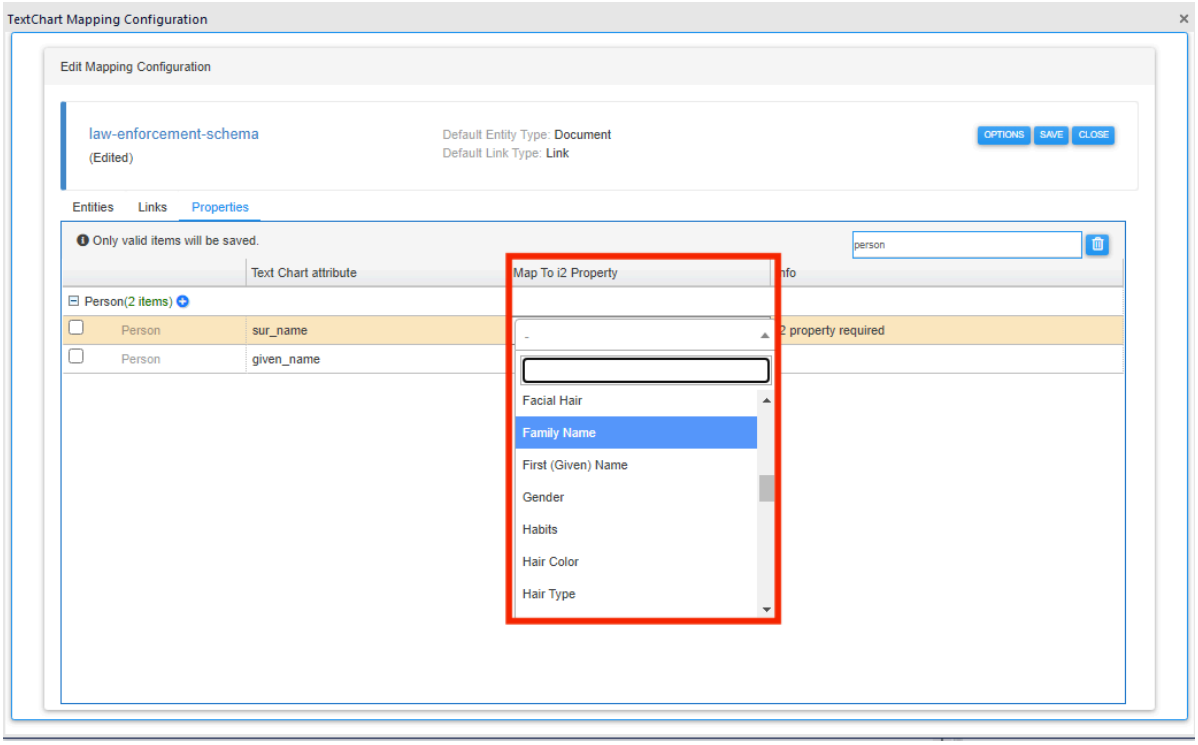
4. In the **Map To i2 Property** column, select the property type of the i2 entity type that you want to map the attribute to.

© N. Harris Computer Corporation



5. Continue this process until you have created mappings for all the property types that you want to see in Analyst's Notebook charts.

© N. Harris Computer Corporation



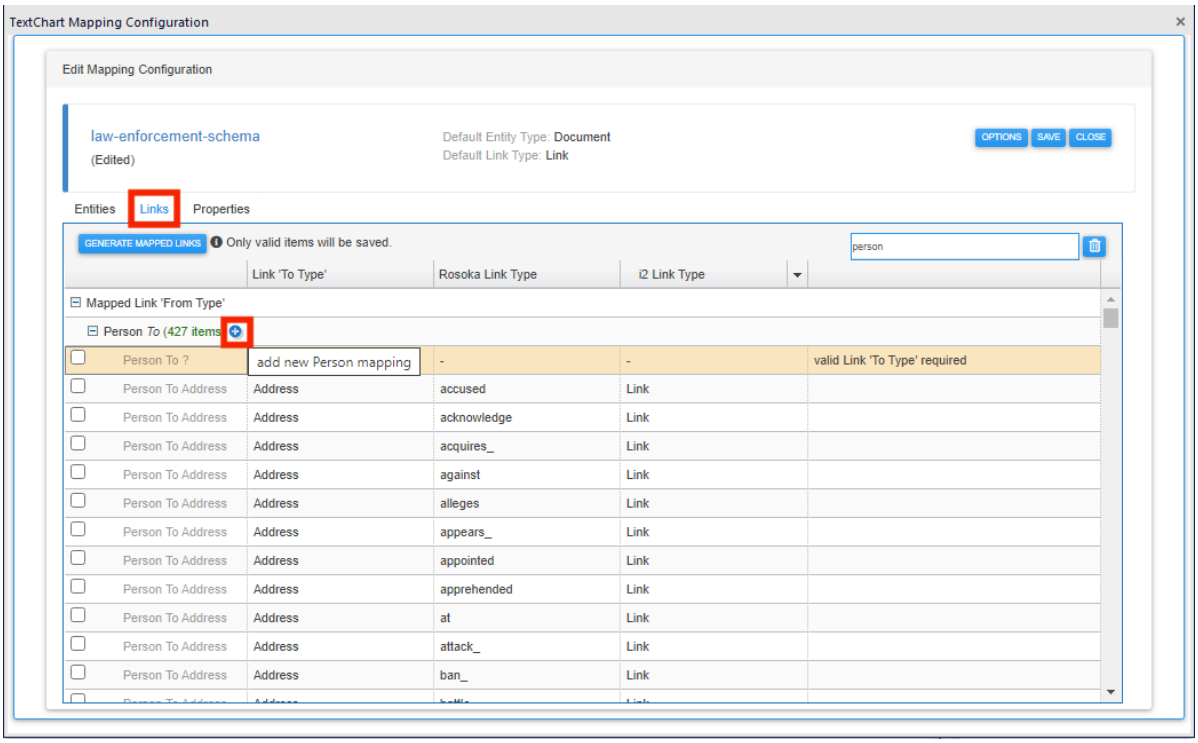
At any point in the mapping process (and always at the end), click **Save** to update the stored version of the mapping configuration.

Mapping link types

You can only create mappings from TextChart link types to i2 Analyze link types after you've created the necessary entity type mappings.

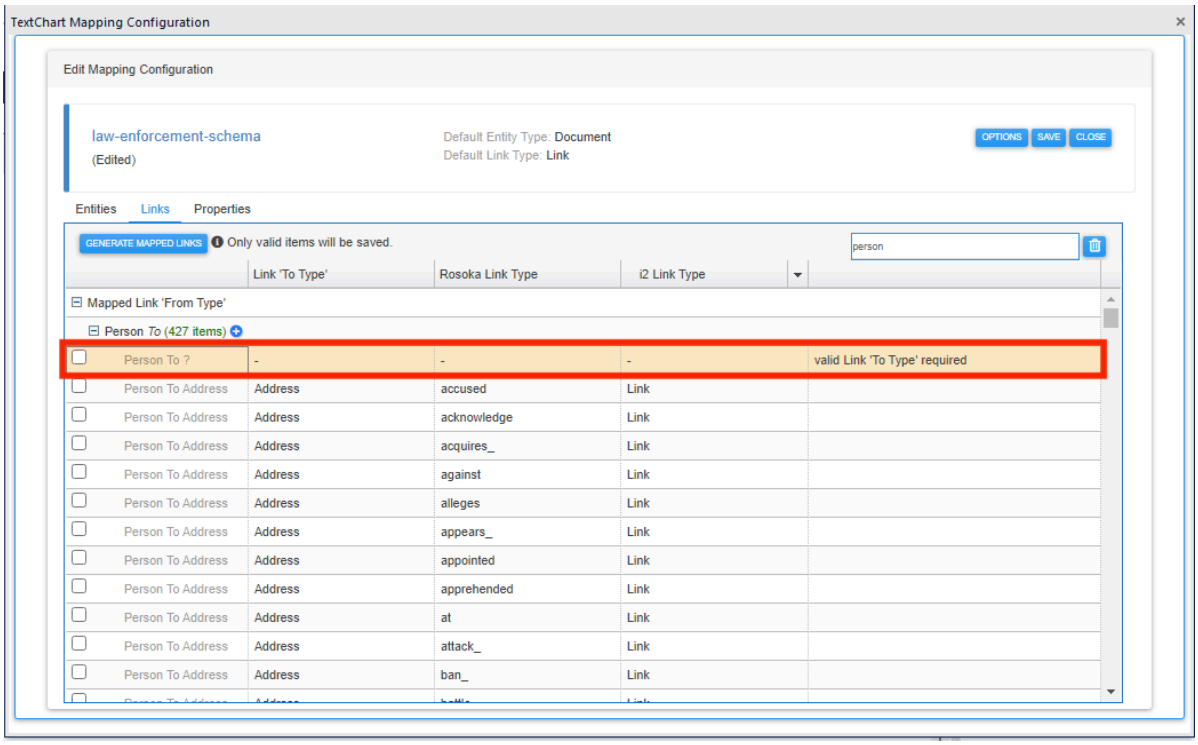
1. Select the **Links** tab in the Mapping Configuration window to display a list of all the entity types that you've mapped.

© N. Harris Computer Corporation



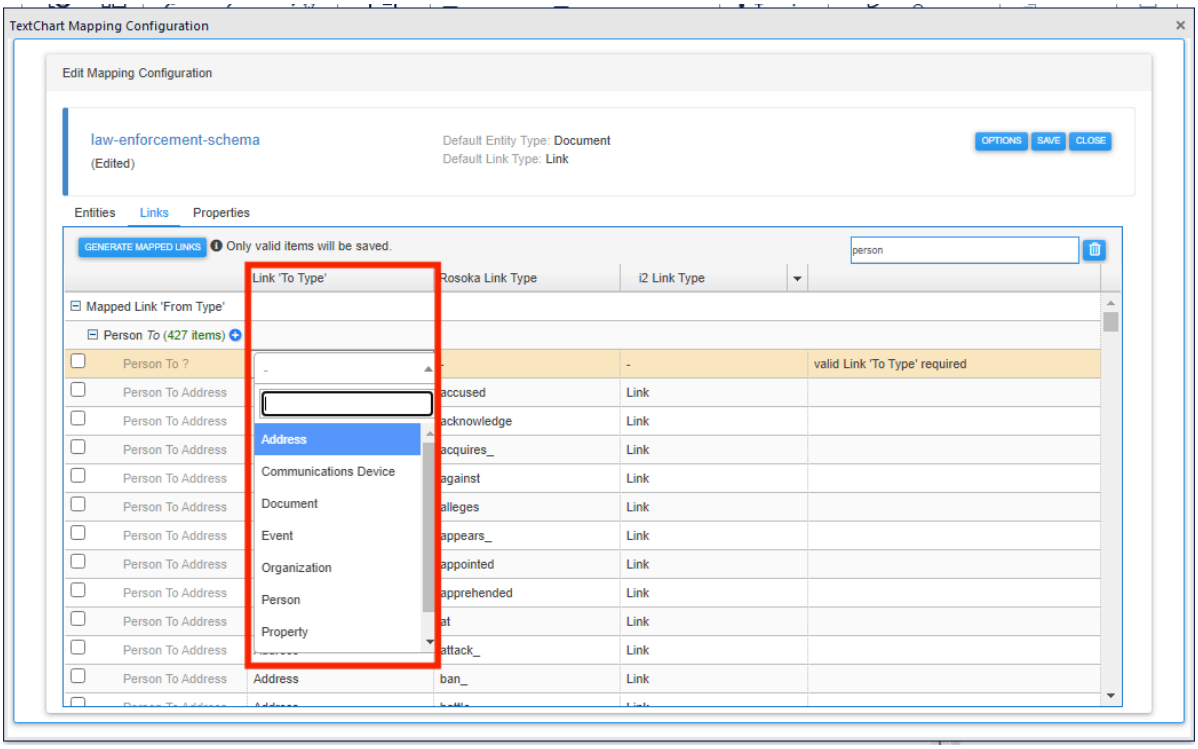
2. Click the "plus" icon next to the name of the entity type that forms the "From" end of the link type that you want to map. A new row appears below the entity type.

© N. Harris Computer Corporation



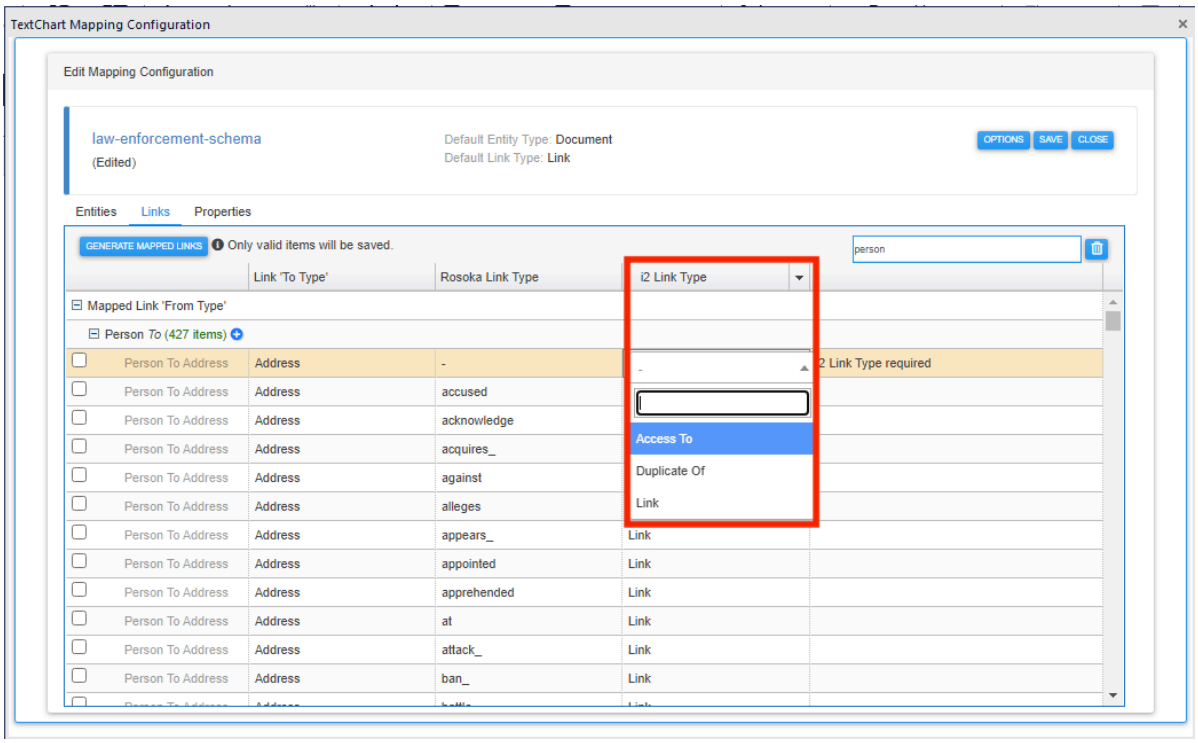
3. In the **Link 'To Type'** column, select the TextChart entity type that forms the "To" end of the link type that you want to map.

© N. Harris Computer Corporation



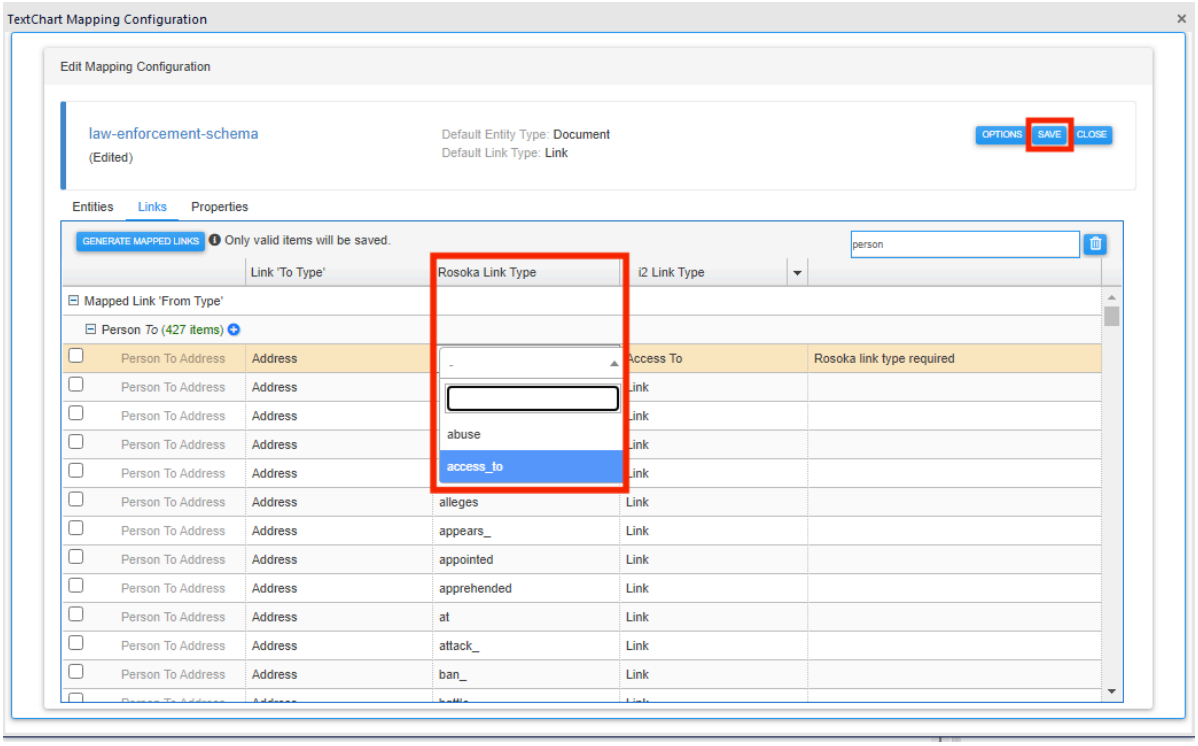
4. In the **i2 Link Type** column, select the i2 Analyze link type that you want to map to.

© N. Harris Computer Corporation



5. Finally, in the **Rosoka Link Type** column, select the TextChart link type that you want to map from, and then click **Save**.

© N. Harris Computer Corporation



Tesseract OCR parser

Tesseract Optical Character Recognition (OCR) parsing may be installed to provide seamless OCR support for many common graphical file formats, including JPEG, TIFF and GIF. Once Tesseract OCR is installed and running, TextChart will seamlessly process supported file formats in the ingestion process via the Tika interface. No further configuration is required.

Installing Tesseract on RHEL

1. Install the EPEL repository if it isn't already available:

   ```
   sudo dnf install epel-release
   ```

2. Install Tesseract:

   ```
   sudo dnf install tesseract
   ```

3. To add language packs, see what's available:

   ```
   dnf search tesseract
   ```

   And then install, for example:

   ```
   sudo dnf install tesseract-langpack-ara
   ```

Installing Tesseract on Ubuntu

1. Download package information from all configured sources:

   ```
   sudo apt update
   ```

© N. Harris Computer Corporation

2. Install Tesseract:

   ```
   sudo apt install tesseract-ocr
   ```

3. To add language packs, see what's available:

   ```
   apt search tesseract-ocr
   ```

   And then install, for example:

   ```
   sudo apt install tesseract-ocr-fra
   ```

Installing Tesseract on Windows

1. Download the latest Tesseract installer from the UB-Mannheim Tesseract repository. Older versions are available at `https://digi.bib.uni-mannheim.de/tesseract/`.
2. Run the installer and follow the prompts. Install Tesseract in the directory suggested by the installer or in a new directory.

   Important: Do not install into an existing directory that contains other files. The uninstaller removes the entire installation directory.
3. During installation, select any additional language packs that you require.
4. Add the Tesseract installation directory to the system `PATH` environment variable so that it is available from the command line.

Installing Tesseract on other platforms

For further details on Tesseract, including installation on Macintosh, see the information at https://cwiki.apache.org/confluence/display/tika/TikaOCR.

© N. Harris Computer Corporation