

Contents

1 Introduction

- 2 Who should read this white paper
- 2 Potential benefits of Text Chart
- 3 Key features of Text Chart
- 11 Technical description

IBM i2 Text Chart

Overcome the challenges associated with the assimilation of unstructured data

Introduction

Analysts, investigators and other workers associated with the collection of information are skilled at processing structured data found in databases and spreadsheets. Discovering vital case information from unstructured data sources such as statements and reports however, is a much more challenging and time consuming process. One can easily overlook important details and lose track of key connections hidden within multiple documents.

IBM® i2® Text Chart is a powerful, intuitive, user-controlled text extraction and visualization application that helps overcome the challenges associated with the assimilation of unstructured data. It quickly transforms text-based information into a structured graphical format, facilitating the production of understandable and valuable materials for analysis purposes.

With Text Chart you can highlight key information within text-based documents, extract, and visually represent those details in an intuitive chart. If you are working with a collection of documents, Text Chart helps you to discover and extract high value information from each document and present your findings as a consolidated view.



Text Chart offers:

- Effective entity-link extraction and visualization of unstructured data.
- Analysis-ready data from text-based content.
- Transformation of information into an intuitive graphical format.
- Visibility of high value information held in unstructured data sources.
- Population of searchable charts and databases without having to re-key information.
- Streamlined input of unstructured free-text information into data repositories.
- Improved user productivity when using IBM® i2® iBase and Text Chart together.

Who should read this white paper

This document has been written for analysts, investigators, researchers and data extractors who wish to learn more about how Text Chart can help them to visualize information contained in unstructured data sources. It presents an overview of the product and the problems it can solve.

Potential benefits of Text Chart

Effective entity-link extraction and visualization of unstructured data

Text Chart helps you identify real world objects including entities, links and their corresponding properties from your unstructured data sources, quickly transforming it into structured, human-validated, analysis-ready data.

Transformation of text-based information into an intuitive chart

Text Chart helps you quickly create a visual representation of what would otherwise be a mass of written reports. The solution is user-controlled, meaning that you decide which information is important and how you want to represent it.

Data abstraction from text based sources

With Text Chart you can extract data directly from your source documents, resulting in a more accurate and consistent extraction process. Your extracted data will be free of the normal errors associated with re-typing or transcribing information.

Discover high value information held in unstructured data sources

As with general sources of information such as intelligence reports, arrest reports and so on, more and more data is now available via e-mail and Web pages. These unstructured data sources contain valuable information that may help you uncover risks, opportunities and make better, more informed decisions. Therefore, as these types of data sources increase, ignoring unstructured data sources is not an option. Using Text Chart, you can tap into unstructured information that would otherwise go unused.

Capture data faster

The assimilation of unstructured data sources is typically difficult and time consuming, requiring huge amounts of human resource. With Text Chart however, the process is optimized, helping to make the data capture process more efficient and thereby increasing user productivity.

Make the most of your data

By delivering powerful text extraction and visualization capabilities, Text Chart supports organizations in their quest to efficiently tap available data sources.

Designed to integrate with other IBM i2 products

The output from Text Chart can be used within other IBM i2 products, accelerating the analytical process. Text Chart output can be used with IBM® i2® Analyst's Notebook®, IBM® i2® Chart Reader, and iBase. Using Text Chart in conjunction with other IBM i2 products enables you to share marked up data with the whole investigating team, helping analysts make more informed decisions.

Key features of Text Chart

Ease of use

Text Chart was designed to be intuitive and quick to learn. This design enables analysts to start marking up documents and creating charts quickly, without having to waste valuable time learning how to use it. Each set of documents and subsequent chart that you create by marking up those documents, are known as a workspace. A workspace can also contain related properties and comments.



Document reading and conversion

You can add many types of document to Text Chart, including Plain Text, Rich Text, HTML, PDF, Microsoft Word, Microsoft Excel, WordPerfect and OpenOffice/StarOffice Write and Calc formats.

Each document you add to a workspace is copied and converted into a format that can be displayed in the Document Pane ready for mark up. This helps to ensure that the integrity of the original document is maintained.



Mark up and visualization

When you have added a document to a workspace, you can begin marking it up. This is the process of highlighting important information to create entities, links and other properties on a chart. In this way, you can create a visual image from your documents.

There are three entity representations: icons (used to represent people, places and objects), theme lines (used to depict an entity over a period of time) and event frames (used to depict events). To show a connection between two or more entities, you can create links. Each type of mark up (entity, link and comment), is represented by a different default color in the Document Pane. You can also define specific markup colors for individual chart items.



Figure 3: Creating an entity involves highlighting the relevant text, and choosing a suitable type (such as a person, place or object) and representation (icon, event frame or theme line)

Multiple documents

Certain circumstances require the ability to mark up several related documents and represent the resulting output as a single chart. Text Chart supports a multiple document environment which enables you to create a single chart using information extracted from several documents.

	REPORT
Int	elligence indicates that <mark>Elymn DUKE</mark> and his <mark>wrife lanet Debble DUKE</mark> are running the
int	emational identity theft network from thei <mark>r home</mark> in <mark>Atlanta: 4770 Burford Highway</mark>
DL	KE nee CARTER is the daughter of the suspect Gary CARTER .
Int	elligence indicates that <mark>Flynn DUKE</mark> , a known <mark>offender, meets regularly</mark> with an ass
pri	son while serving time for <mark>fraud. William KNAPPENBURGER</mark> was convicted for <mark>mone</mark>
off	ences in 1999.
Int Blo	elligence indicates that <mark>Flynn DUKE</mark> and his wife <mark>Janet Debble DUKE</mark> have a female <mark>India</mark> who <mark>drinks in Tricksters Bar</mark> with <mark>Samuel WAUKER</mark> . She drives a red Honda Ac ER and is described as a young white female with fair hair.
of	ficer research – the red Honda driven by <mark>Blandie</mark> is registered to <mark>Sarah PIKE,</mark> born 0
19	82 in Memphis. Her registered address is 81/78 Broadway, Montgomery Village, MD
nu	mber of 260 321 9056 is associated with this address.

 $\it Figure~4:$ Each workspace document is displayed as a separate page in the Document Pane, indicated by a tab

References to the original source documents are maintained, meaning that sources are traceable.

Template design

Each Text Chart workspace is based on a template which determines the type of information that you can select when you create items on a chart. When you create a new workspace you can base it on a template so that it displays information in a way that is consistent with other workspaces in your organization.

A predefined Standard template, which is suitable for most types of general investigation, enables you to start using Text Chart immediately. If your organization has more specific requirements however, you can use Text Chart Designer to design and manage your own templates. Text Chart Designer enables administrators to define many workspace properties, for example:

- Entity and link type definitions, including semantic properties.
- · Attribute definitions, including semantic properties.
- Chart properties such as the grading system and date and time formats.

mplate - Standard						
Account	hart Properties	_			2 🗙	
- Dubling Cash Cash CallPhone Conscolity	Definitions	Pabite: Standard		Show Small to	ns	
Courses Courses	r dit Jenty Type Nane: Perak Na Perak Na Perak Jese Cike: Des Color: Jesentik Type Person	Kensel	Painties: Common Retribles Communities Communities Nonno Communities Nonno Sociality Nonno Sociality Nonno Non Non			
			~	[Cont] [

Templates can be customized based on an organization's unique needs, adapted from the Standard template, or based on a template or chart from another IBM i2 application.

Property values

After you have created an entity or link, you can add property information to it in the form of attribute values. For example, you can record important details such as a person's nationality, age or distinguishing marks; the color of a vehicle and its license plate; or a description about an incident including reliability grades.



		Name		Value	^
	Ð	Last Name		CARTER	
٠	Θ	First Name	3	DEBBIE Deborah Debs	
		Cobbie CART	ER 🔿	DEBBIE	
		C Deborah		Deborah	
		C Debs	-	Debs	
				<click inferred="" insert="" to="" value=""></click>	

The attribute values that you can add to an entity or link depend on the type you have chosen. Each entity and link type has specific attribute slots associated with it.

You can add an attribute value by marking up text or, if the value is implied, you can add an inferred attribute value (see Inferred property values and links, next). Other attribute values are default values that are specified in the template. Attribute values can also be derived automatically from marked up text when an entity is created, using semantic type information that is specified in the template. The values that are derived depend on the chosen entity type and whether defining values are included in the marked up text.



Figure 5: Depending on the selected entity type and marked up text, attribute values can be derived automatically when an entity is created

Each attribute slot and value associated with a chart item is listed in the Properties Window. Different marker symbols indicate whether a value is a mark up, inferred, default or derived. Attribute slots can also contain multiple values, allowing you to record the data that is associated with a particular slot.

Inferred property values and links

At certain times when you are marking up a document, you may come across relationships and other information that are implied in the text. You may infer from a statement or other document that a particular individual is a suspect or victim, even though this information is not explicitly stated, or you may want to create a link between entities which does not directly relate to a document mark up. In these cases, you can add an inferred attribute value or create an inferred link without marking up text in a document.



Figure 6: Adding an inferred attribute value is a simple case of typing the required text in an appropriate slot in the Properties Window

The option to allow inferred values is controlled by settings in the template, and can be enabled or disabled using Text Chart Designer.

Grading system

Grades are used to classify information, typically providing evaluations about quality, reliability and confidentiality. A grading system which meets your organization's requirements can be set up in the template, enabling the consistent use of grades in your workspaces.

Text Chart allows up to three categories of grade to be assigned. You can choose whether to assign grades to chart items or to the individual mark ups in a document, or to both.



Figure 7: You can mark up grading evaluations in a document and assign them to appropriate chart items. Chart items automatically inherit source information from associated documents

You can also enter source information about each document in your workspace by assigning a source type to record the origin of the information (witness, informant, officer and so on) and a source reference to provide a link back to the original source. This source information is then inherited automatically by chart items and mark ups associated with the document.

Powerful search

A powerful search tool helps you to locate specific text, identical and similar entities, related mark ups and related chart items within your documents and chart, facilitating the mark up process. Synonym, spelled like and sounds like fuzzy search options allow you to find results that are similar to a search term as well as exact matches.

There are three different searches:

- Find in Documents enables you to find and mark up each reference to an entity in your workspace documents, using the results to create a new chart item or add data to an existing item.
- Find in Chart enables you to search for text in the labels, property values and associated mark ups of entities and links on the chart.
- Find target items enables you to search your workspace documents and chart for significant items defined on a target list, using the results to create a new chart item or add data to an existing item. This type of search can be triggered automatically when a document is added to the workspace.

	Doble CARTER Define Gary CARTER Or Janet Debble DURE
Find Result	lts ?
Target Items:	4 matches found for 'Gary CARTER':
Using Debbie C	CARTER 3 Group by Document •
	Junet DUKE Intelligence JunetH Intelligence indicates that Garry CARTER has a daughter called Janet
	O Cruste new chart Rem
	O Create new chart Rem O Add occurrence(s) to checked chart Rems: Existing Chart Rems

Figure 8: The search results are organized as a tree view. You can specify which matches to mark up and how to apply them to the chart. Interaction with both the Document and Chart Panes allows you to zoom to matches in a document or items on the chart

Association and timeline charts

Text Chart can be used to create association charts, timeline charts, or charts that exhibit a combination of both association and timeline properties.



Association charts (also known as network charts) show the relationships between entities to illustrate how they are interconnected. For example, associations between individuals, transactions between accounts, calls between telephone numbers and so on. This type of chart does not tend to contain items which have date and time.

Timeline charts show a chronology of events over a period of time, such as a series of meetings occurring over several days, or a set of transactions occurring over a period of time. This type of chart shows theme lines and contains items which have date and time. You can demonstrate the flow of time across a chart by turning on the time bar.



Chart layouts

Layouts provide a quick way to automatically rearrange the entities and links on a chart, improving the appearance and making it easier to interpret. The layout you choose depends on the type of chart you are working with.



The Minimize Crossed Links layout rearranges your chart to minimize crossed links between entities. This layout is useful when you are working with an association chart, enabling you to arrange the entities into the clearest layout possible.



If you are working with a timeline chart you can apply the Proportional layout. This layout rearranges charts so that items with date and time are spaced apart as they are in relation to each other in time.

Chart legend

Information about a workspace can be displayed on a chart in the form of a legend. You can use a legend to display fixed text, such as a key of entity and link types, or placeholder fields, such as Title, Author and Comments, that are populated automatically from information that you enter as workspace summary properties. When you print a chart, the legend is also printed, providing a record of important information.



Comments

Comments can be added to a workspace to annotate sections of text in a document. This is useful if you need to draw attention to a particular area of the document, or indicate where information is incorrect or needs further clarification. Adding a comment is a simple case of marking up the text that you want to annotate and then entering the comment text in the Comments Window.



Figure 9: Moving your mouse cursor over a comment mark up in the Document Pane causes the comment text to display as a tooltip

Template report

You can generate a report to provide information about the template on which your workspace is based. The report allows you to visualize the template contents, giving details about each entity type, link type and attribute that is defined, as well as which attributes are assigned to which entity and link types as slots.

The report can be generated from both Text Chart and Text Chart Designer.

Contents of the template Standard: Interts of template standardde standa	View Favorites Tor	ols Help			
Interts of the template Standard: tity Types https:	0 · 🖻 🖬 🔞	🔎 Search 🔥 Pavorites	🛛 🙆 · 🖓 🕅 ·	D 🕰 🖪	
tity Types k Types https://www.internationality.com/ https://www.internationality.com/ https://www.internationality.com/ Account Number None Account State None Stat Code None Stat Code None	ntents of the templa	ate Standard:			
Corresponding Chart Rem Default Value Stot Markers title Slot Markers Interface Interface Interface title Types Account Interface Stot Markers Account Person Stot Markers Interface Account Person Stot Markers Interface Account Name One Interface Interface Account Name None Interface Interface Start Code Grade1 Interface Interface Start Code Grade2 Interface Interface					
h. Lydon histeles histeles Site Markers te: Inferred attribute values are enabled for this template.	ty Types				
	ibutes				
tity Types Account Account Number None Account Number None Account Type None Soft Code None None None None None None None Non	ibute Slot Markers				
Account Stot Markers Account Preserver Account Number None Account Number None Account Type None Account Number None Account Number None Account Type None Account Type None Account Type None Account Type None Bank Name None Static Code Static Code	. Informed attack the second				
Account Corresponding Chart Rem Pendule Default Value Stot Markers Account Number None Account Number None Mone Mone Account Number None Mone Mone Mone Bank Name None Mone Mone Start Code None Mone Mone Mone Start Code None Mone	: interred attribute val	ues are enabled for this ti	emplace.		
Account Activate Corresponding Chart Item Proverty Account Number Account Name Account Name None Bank Name None Sant Code None Code Sant Code None Code Sant Code None Code Sant Code San	ity Types				
Account Account Corresponding Chart Item Default Value Stot Markers A Account Number None Account Number None Account Number None Account Number None A Account Number None Account Number None Account Number None Account Number None Bank Name None Account Number None Account Number None Account Number None Sand Cade None Account Number None Account Number None Account Number None Sand Cade None Account Number None Account Number None Account Number None Sand Cade None Account Number None Account Number None Account Number None Account Number None Sand Cade None Account Number None Account Number None Account Number None Account Number None Sand Cade None Account Number None Account Numer None Account Numer None					
Vertoute Corresponding Chart Name Default Value Stot Markers 4 Account Number None <					
Account Number None Account Name None Bank Name None Sant Code None Sant Code None Code I Description Grade1 Grade2 Grade2	Account				
A Account Name • Account Type • Bank Name • Bank Name • Sant Code • Rescription • Grade 1 • Grade 2	Account	Corresponding Chart Item Property	Default Value	Slot Markers	
Account Type None Bank Name None Sort Code None Description Grade1 Grade2	Account Account	Corresponding Chart Item Property None	Default Value	Stot Markera	
Bank Name None X Sart Code None 3 Description - 6 Grade 1 Grade 2 4 Grade 2 Grade 2	Account Atribute Account Number Account Number	Corresponding Chart item Property None	Default Value	Slot Markers	
X Sart Code None 1 Description Description c Grade 1 Grade 1 x Grade 2 Grade 2	Account Atribute Account Number Account Number Account Name Account Type	Corresponding Chart Item Property None None None	Default Value	Slot Markers	
I Description I Grade 1 Grade 2 Grade 2	Account Atribute Account Number Account Number Account Name Bank Name	Corresponding Chart Ben Property None None None None	Default Value	Stot Markers	
* <u>Grade 1</u> Grade 1 - ** <u>Grade 2</u> Grade 2 -	Account Aeribute Account Number Account Number Account Name Bank Name Sort Code	Corresponding Chart Item Property None None None None	Default Value	Slot Markers	
I Grade 2 Grade2 -	Account	Corresponding Chart Bees Property None None None None Description	Default Value	Slot Markers	
	Account	Corresponding Chart Bean Property None None None None Description Grade1	Default Value	Slot Markers	
Grade 3 Grade3 -	Account Aerout Account Number Account Number Account Name Account Name Bank Name Sort Code Description Grade 1 Grade 2	Corresponding Chart Ress Presenty None None None Description Grade1 Grade2	Default Value	Stot Markers	

Figure 10: The report is generated in your Web browser, where you can print it, send it by e-mail, save it, or copy and paste it into another application

Printing

When printing a workspace you can finely control the printing setup. You can choose which parts of your workspace to print (the chart only, the marked up documents only, or the chart plus the documents), the printer, paper size and orientation, what scale your chart should be printed at, the position of the legend and so on.

IBM Software White Paper

When specifying page printing options for the chart, a print preview based on current page boundary, page orientation and scale settings is displayed in the Chart Pane.

The report is generated in your Web browser, where you can print it, send it by e-mail, save it, or copy and paste it into another application.



Figure 11: The Print Preview displayed in the Chart Pane reflects the current margin, page boundary and scale selection. Page boundaries are displayed as dashed lines so that you can clearly see how the chart will be printed

Workspace customization

You can customize a workspace by repositioning, resizing, auto hiding, or closing the Palette Bar and the Properties Window, or by adjusting the relative sizes of the Chart and Document panes. Different layouts are more suited to different ways of working. For example, if you do not tend to use the Palette Bar when creating entities, you may decide to close it or set it to auto hide; if you are using the Properties Window to add many attribute values, you may decide to dock it to the side of the application window so that each attribute slot and value is visible. You should be able to find a layout that is best suited to the way you work.



Figure 12: When a window is set to auto hide, it is minimized to a tab in the application window. Dragging text onto the tab or clicking on it, causes the window to pop up temporarily

Compatible with other IBM i2 products

Text Chart can be used alongside other IBM i2 products including Analyst's Notebook and iBase. Text Chart supports semantic types — this is a way of tagging information to give it a uniform meaning that is understood across all IBM i2 products.

When used as a standalone application, Text Chart applies semantic technology to derive attribute values from key information contained in marked up text. If you intend to use your workspaces in conjunction with other IBM i2 products the potential benefits from semantic technology are even greater.

This helps enable you to:

- · Perform more intelligent searching, matching and analysis.
- Analyze data and combine data more conveniently.
- Spend less time finding data and more time analyzing results.
- Extract more value from existing data.

Using Text Chart with Analyst's Notebook

You can open Text Chart workspaces with Analyst's Notebook, enabling its full range of powerful visual, investigative and analysis tools for your workspace data. The marked up workspace documents are displayed as OLE objects which you can open and view, and markups are stored as cards. If your chart items are semantically typed, you can use Analyst's Notebook's smart matching functionality, to help overcome different recordings of the same information, accidental spelling mistakes and different formats of the same information. This can mean that previously overlooked key associations are drawn to your attention.



Figure 13: Workspace documents are displayed as packaged OLE objects which can be opened and viewed in your Web browser application. An item's associated mark ups from Text Chart are stored as property cards

Text Chart workspaces opened in Chart Reader are also displayed in a similar way to that described above.

Using Text Chart with iBase

You can transfer data between Text Chart and iBase. Using Text Chart in conjunction with iBase, you can push unstructured text stored in iBase into Text Chart, extract and visualize the important information, then populate your iBase database with the structured results. There is no need to re-key information, giving you complete control over your data.

There are two main methods of data transfer:

- Using a live connection to an iBase database you can set up associations between items on a chart and records in a database, enabling the transfer of information between Text Chart and iBase as you mark up. You can add new records to the database based on the properties of items on a chart, and create new chart items using information from the database, updating and refreshing the data as required.
- Using a visualization loader you can load a Text Chart workspace into an iBase database. During the load process, new entity and link records are created automatically; you are prompted to intervene only if data from the workspace matches an existing record in the database, or if some mandatory data is missing. If you have used semantic typing, this information is used to help with alignment of data between the workspace and the iBase database. This is particularly useful if changes have been made to the database schema since the Text Chart workspace was created.



Figure 14: The loader allows you to create workspaces in Text Chart and subsequently transfer the information to a database The live connection allows you to transfer information between a workspace and the database in real time

Technical description

Product architecture

Text Chart is a standalone, out of-the-box desktop product that consists of two applications:

- Text Chart.
- Text Chart Designer.

Text Chart can be used in conjunction with Analyst's Notebook, Chart Reader, iBase and other i2 products.

What documentation is provided?

To help you get the most from your investment, Text Chart is provided with a comprehensive set of user and quick start guides, online help and release notes. The tutorial-style user guide is supported by example workspaces and documents that illustrate the features you can use in your workplace.

The documentation provided with Text Chart is available electronically in Adobe PDF format.

Is Text Chart available in languages other than English?

Text Chart is an internationalized application that can be localized within the supported language environments. It is written in US English and the support language environments are US and Western European, Central European, Turkic, and Cyrillic operating systems.

Contact your supplier for further details regarding available local language versions.

How does Text Chart integrate with other IBM i2 products?

Text Chart can be used in conjunction with suitable, compatible versions of many other IBM i2 products, including Analyst's Notebook, iBase and Chart Reader.

Can I upgrade from other IBM i2 products?

Text Chart is only supported when used with compatible versions of IBM i2 products. If you are using a combination of previous versions of these products you must upgrade them all at the same time in order to maintain compatibility.

Implementation and training

IBM offers standard training courses for Text Chart both for analysts and template designers. These are designed to help your staff get immediate time-saving and analytical benefits from your new system. Consulting options are also available to support your deployment.

For more information

To learn more about IBM i2 Text Chart, please contact your IBM representative, or visit: ibm.com/i2software

To learn more about all of the IBM Smarter Cities solutions, visit: ibm.com/smartercities



© Copyright IBM Corporation 2012

IBM Corporation Software Group Route 100 Somers, NY 10589 U.S.A.

Produced in the United States of America

November 2012

i2, Analyst's Notebook, COPLINK, IBM, the IBM logo and ibm.com are trademarks of International Business Machines Corporation in the United States, other countries or both. If these and other IBM trademarked terms are marked on their first occurrence in this information with a trademark symbol (® or TM), these symbols indicate U.S. registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. Other product, company or service names may be trademarks or service marks of others. A current list of IBM trademarks is available at "Copyright and trademark information" at: ibm.com/legal/copytrade.shtml

The content in this document (including currency OR pricing references which exclude applicable taxes) is current as of the initial date of publication and may be changed by IBM at any time. Not all offerings are available in every country in which IBM operates.

The performance data discussed herein is presented as derived under specific operating conditions. Actual results may vary. THE INFORMATION IN THIS DOCUMENT IS PROVIDED "AS IS" WITHOUT ANY WARRANTY, EXPRESS OR IMPLIED, INCLUDING WITHOUT ANY WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND ANY WARRANTY OR CONDITION OF NONINFRINGEMENT. IBM products are warranted according to the terms and conditions of the agreements under which they are provided.



Please Recycle