

Red paper

LexisNexis Bridger Insight XG Matching Algorithm

A guide to reducing false positives

Released July 2011

LexisNexis® Bridger Insight™ XG is designed to help organisations of all sizes, across various industries, comply with regulations.



The idea of watchlist checking sounds simple enough. Regulators ask you to make sure you are not doing business with entities identified on watchlists. Easy, right?

When you consider how many people and organisations you do business with, how many lists there are to check against, and how often these lists change - the job becomes daunting. Bridger Insight XG can help alleviate this burden by providing watchlist searching. By allowing you to influence the potential matches that are reported, Bridger Insight XG can help you meet your compliance needs based on your organisation's level of risk.

Executive summary

When you consider how many people and organisations you do business with, how many lists there are to check against, and how often those lists change – the job becomes daunting. If your organisation has one million customers, you are checking three watchlists, and those lists are typically updated two times a month, that is six million comparisons per month and 24 million per year. This figure multiplies greatly when you count the number of AKAs (also known as), addresses, and IDs, which can be included in watchlist entries. You can quickly see that with the number of comparisons necessary, automation is key to completing this process in a timely manner.

LexisNexis® Bridger Insight™ XG is designed to help organisations of all sizes, across various industries, comply with regulations such as OFAC, USA PATRIOT Act, and the EU Third Money Laundering Directive. With its configurable watchlist searching capabilities, this tool helps you mitigate compliance risk while giving you some control over how comparisons are performed. Its sophisticated, articulate algorithm operates using fuzzy logic to help you find more true matches and fewer false positives. Bridger Insight XG can help you reduce watchlist checking times and help reduce associated risk.

When you rely on automation to perform watchlist comparisons for you, you need to understand how it works. This technical white paper describes some of the matching processes.

Watchlist checking

The OFAC and the USA PATRIOT Act, and EU Third Money Laundering Directive prohibit financial institutions from conducting business with certain entities. To abide by relevant laws and regulations, organisations check their input information against watchlists to alert them of possible prohibited transactions.

Algorithm overview

The Bridger Insight XG matching algorithm is formulated to find potential matches despite irregularities, such as common entry errors, while limiting the number of false positives reported. After identifying candidate records, Bridger Insight XG further analyses data elements to generate a confidence score and create alerts.

You can choose to have Bridger Insight XG find candidate records when specific data elements match allowing you to find more connections between input data and watchlists. You can also set rules to prevent Bridger Insight XG from reporting alerts you pre-determine to be false positives using settings called Automatic False Positive rules.

In the rest of this section, we discuss some of the fundamentals of searching to give background and common vocabulary as we explore the Bridger Insight XG matching algorithm.

Search Entities

In Bridger Insight XG, you can input these subjects to compare against watchlists:

- People
- Businesses
- Unknown

Bridger Insight XG compares your input subjects to watchlist entries.

The watchlist entries may be any of the following types:

- People
- Businesses
- Countries
- Vessels

When referring to search subjects or watchlist entries generically, we may call them “entities.”

List Types

You can search for entities in these lists:

- Standard watchlists – Provided with a Bridger Insight XG annual license. LexisNexis® Risk Solutions maintains these lists of information from U.S. and international agencies. Some lists contain only country data. Separating people/business and country data into different lists helps Bridger Insight XG improve matching accuracy.
- Custom watchlists – Created by your organisation by importing information or adding records individually. These lists can contain information unique to your organisation. Many Bridger Insight XG clients use this capability to host FinCEN4 data.
- Accept lists – Created by your organisation over time from within Bridger Insight XG. Users at your organisation add false positive matches to these lists. When included in searches, the lists help suppress repeated false positive alerts. Records must include a sufficient address or a unique ID for users to add them to accept lists.
- PEP data sets – Purchased from third-party providers as an optional addition to Bridger Insight XG. These large data sets include information on politically exposed persons (PEP) worldwide.

When we refer to these different lists as a group, we typically call them “watchlists.”

A key feature of Bridger Insight XG is your ability to control several criteria used to report alerts.

Data Structure

The entity records you need to search may exist in different databases, files, and transaction formats at your organisation. The structure of this information may be different in every instance. Bridger Insight XG can handle data in a number of different formats. Bridger Insight XG can process data that is parsed or unparsed. It can also process data in common file formats that contain a combination of parsed and unparsed content.

Search Criteria

A key feature of Bridger Insight XG is your ability to control several criteria used to report alerts. An abundance of watchlist information is not unique and can result in many false positive matches. Bridger Insight XG lets you control search settings based on your risk tolerance to help reduce false positives.

Bridger Insight XG reports potential matches as alerts based on the criteria that you have specified in bundled search parameters, called a “predefined search” depending on your Bridger Insight XG delivery method. The criteria includes, for example, a minimum confidence score. Bridger Insight XG reports alerts against selected watchlists if the confidence score of the potential match is above a minimum score you have set for each list. This feature lets you control how closely information must match before Bridger Insight XG reports it as an alert.

With parsed files, you can also configure match options to have Bridger Insight XG identify candidate records using data elements other than name. Looking for matches using address, ID, or phone number can decrease the speed of the search and increase the number of potential matches reported, but it can help you find links to watchlist entities that may not be clear based on name-only searches. This feature might be helpful if you want to apply additional scrutiny to high risk customers.

You can also instruct Bridger Insight XG to not report matches, not flag alerts, or set record statuses to “Automatic False Positive” when selected data elements do not match in parsed files. You can apply these Automatic False Positive rules to address, date of birth, gender, ID number, phone number, and entity type data elements.

Using Bridger Insight XG, you can define your own riskbased approach to watchlist checking with these search parameters to max-

imise the number of true matches found while minimising the false positives reported. Configuring the matching algorithm to return only alerts that exceed your risk tolerance may lighten your compliance processing load. The quality of potential matches you allow Bridger Insight XG to report (or you choose to disregard) is a business decision you can make based on your organisation’s risk tolerance.

Search Process

Depending on the search and input information, Bridger Insight XG performs different types of checks against the data. For example, unparsed data is searched using different methods than parsed data, and people searches undergo different processes than business searches. That said, for the most part, Bridger Insight XG follows the same general process when it performs searches. These are the general steps.

1. Find candidate records
2. Analyse the candidate records to generate a confidence score
3. Check accept lists included in the search to remove candidate records
4. Apply automatic false positive rules to further reduce candidate records
5. Report candidate records as alerts

Next we more closely examine some of the logic and scoring processes used during these steps.

Entity Scoring

During each search step, Bridger Insight XG performs a great number of comparisons and calculations to determine the final confidence score. Bridger Insight XG can perform searches on name alone. However, one of the great strengths of Bridger Insight XG is that it can score on more data elements than name alone if the data is correctly parsed and provided. Names are generally not unique. By comparing additional data elements, Bridger Insight XG further assists you in determining the strength of a potential match. Depending on the type of search and input information, Bridger Insight XG can compare address, ID number, date of birth, gender, and phone number data elements. This allows Bridger Insight XG to provide entity scoring, which can help reduce false positives and generate more robust scoring to help you manage your risk exposure.

Fuzzy Logic

Bridger Insight XG features fuzzy logic search comparisons. Fuzzy logic is a type of logic that allows for values other than true or false. With fuzzy logic multiple values are possible, and this allows Bridger Insight XG to better reflect small variations in data. It also requires that Bridger Insight XG make many, many comparisons as it works through calculations to determine the confidence score for an input record. The following are some of the considerations Bridger Insight XG takes into account during its fuzzy logic comparisons.

- Abbreviations
- Anglicised words
- Extra Words
- Foreign Transliterations

- Initials
- Intervening Words (business only)
- Missing Double Letters
- Misspellings
- Nicknames
- Run-together Names
- Titles
- Typographical Errors
- Word Order (Business Only)
- Word Variation

The following two tables show how these fuzzy logic techniques may identify matches with names and addresses.

Name Match Examples

Naming Consideration	Input Name	Watchlist Entity Name
Abbreviations	Shoes Inc.	Shoes Incorporated
Anglicised Words	Michael Munez	Miguel Munez
Extra Words	Red River Rock Store	Red Rock Store
Foreign Transliterations	Mohammed Jazini	Muhamed Jazini
Initials	T. Arthur Kane	Theodore Arthur Kane
Intervening Words	IRI Shipping and Trading Lines	IRI Shipping Lines

Name Match Examples

Naming Consideration	Input Name	Watchlist Entity Name
Missing Double Letters	Helga Bekemelem	Helga Bekkemellem
Misspellings	Collene McMurphy	Colleen McMurphy
Nicknames	Alex Ivanov	Alexander Ivanov
Run-together Names	Kaylee Miller	Kay Lee Miller
Titles	Javier Labrun	Dr. Javier Labrun
Typographical Errors	Shakevisin Choqdhury	Shakevisinh Chowdhury
Word Order	Ti Wu	Wu Ti
Word Variation	Industry Group	Industries Group

Address and Phone Match Examples

Data Consideration	Your Customer Data	Watchlist Data
Abbreviations	PO Box 123	Post Office Box 123
Misspellings	Rural Root 1	Rural Route 1
Run-together Numbers	5556667777	555-666-7777
Typographical Errors	Mississippi Ave	Mississippi Ave
Extra Words	123 Main St Apt 89	123 Main St

Unicode and Normalisation

Why is Unicode important? Unicode allows users to compare entity data in all supported alphabets. It also allows Bridger Insight XG to normalise comparison information to standardise input text, so it can be more reliably and accurately compared. Bridger Insight XG performs normalisation with no appreciable loss of processing speed.

For example with normalisation, spelling variations in the names Chloë and Chloé would both become CHLOE, which would allow for more accurate matching.

Normalisation also works with alphabets other than Latin. For example with Arabic normalisation, the initial, middle, and end forms of letters are all converted to the isolated form. So the name “Jamal”— becomes with normalisation.

While normalisation is used for comparison purposes with Bridger Insight XG, reported results are not normalised. Watchlist entries and input data are provided in the results as they were originally entered.

Missing/Initialised Last Name Examples

Watchlist Entity	Thomas Jeffrey Madison
Likely Match	Thomas Madison
Unlikely/Low Match	Thomas Jeffrey, Thomas M.

Scoring Results

All potential matches are assigned a confidence score. The confidence score is not a percentage, but an indicator of how closely your input information matches a watchlist entity. This score can range from 0 (low) to 100 (high).

Score Generalisations

Score Range	Description
100	Exact Match of Data Elements Considered
99-95	Highly Likely Match
94-90	Very Good Match
89-85	Good Match
84-80	Less Likely Match
79-75	Weak Match
74-0	Unlikely Match

Minimum score example

You set the watchlist minimum score to 80. When you search your input information, Bridger Insight XG finds potential matches with confidence scores of 65, 72, 77, 89, 95. Bridger Insight XG only reports the two matches with scores above 80.

Nicknames and Word Variations

To identify name variations, Bridger Insight XG compares the input to an internal file of nicknames, abbreviations, and anglicised words. Currently, the extensive file lists mostly Latin-based variations. This file is dynamic, however, allowing LexisNexis Risk Solutions to continually add name variations.

False Positive Reduction

The Bridger Insight XG matching algorithm can eliminate a large number of false positives from being reported as potential matches. False positives are potential matches in which the input information is similar to the data contained in a watchlist, but the input entity is not actually the watchlist entity.

When searching input information against watchlists, Bridger Insight XG factors in frequent causes of false positives such as missing or initialised last names and common words. Like name variations, the matching algorithm compares the input information against an internal file of common words. Bridger Insight XG weights common words lower than more unique words.

Common Term Examples

Watchlist Entity	Bird Bath Factory
Likely Match	Bird Bath Company
Unlikely/Low Match	Bubble Bath Factory

A score of 100, for example, does not mean that your input entity is the watchlist entity. It means some of the input information matches watchlist entity information, and you may want to further examine the alert to determine next steps to take if any. Only potential matches that meet the minimum score threshold defined by you for the watchlist are reported. You set this minimum score threshold for each watchlist searched.

Efficiently address compliance requirements and protect your organisation's reputation.



Summary

The quantity and complexity of data encountered during watchlist checking makes a powerful search tool vitally important in lessening your manual workload and helping ensure relevant matches are found. This need is compounded by the legal, financial, and reputational consequences your organisation may experience if you are found doing business, even unwittingly, with barred entities. Bridger Insight XG can help alleviate this burden by providing watchlist searching that features a matching algorithm that systematically resolves many of these data issues.

By allowing you to influence the potential matches that are reported, Bridger Insight XG can be used to help you meet your compliance needs based on your organisation's level of risk. This configurability may save you time and resources while allowing you to efficiently address compliance requirements and protect your organisation's reputation.

Trust LexisNexis to protect your business

LexisNexis has a world-class reputation for providing professional firms with critical business tools. For over 30 years we have been pioneers in risk management and intelligence.

Our solutions are used internationally by financial services, legal and accountancy firms and blue chip multinational companies, including the world's top 5 banks, to enhance business decision making, reduce the cost of compliance, fulfil regulatory requirements and prevent money laundering.

Access all the UK and global news, company and individual information, sanctions, PEP and watchlists you need for cost-effective and efficient client and third party screening, enhanced due diligence and media monitoring.

Get in touch

To find out more about LexisNexis Bridger Insight and how LexisNexis can help your business: visit us online at www.lexisnexis.nl write us a email servicedesk@lexisnexis.nl or call an expert at +31 (0)20 485 34 56

LexisNexis Benelux, Amsterdam



T +31 (0)20 485 34 56
E servicedesk@lexisnexis.nl
W www.lexisnexis.nl