An Oracle Best Practice Guide
August 2011

Oracle Watchlist Screening: Technology Guide

ORACLE

1. Document Purpose

This document provides an overview of the capabilities and technical characteristics delivered by Oracle Watchlist Screening, a risk and compliance screening application. Oracle Watchlist Screening provides the platform for a growing portfolio of screening solutions tailored for specific markets, including

- Financial services providers, including retail and investment banking, insurance, and trade finance
- Third-party logistics providers and supply chain operators
- Mobile operators offering mobile money transfer services
- Any organization engaged in the export of materials and/or finished goods

This document is intended to give IT staff an understanding of how Oracle Watchlist Screening will integrate into their environment, and answer frequently asked questions, such as

- What hardware and software are required?
- What connectivity to internal and external sources is required?
- In what format does our customer data need to be provided?
- What disaster recovery considerations should we make?

2. Oracle Watchlist Screening Overview

The primary function of Oracle Watchlist Screening is to provide an end-to-end process for matching any individual or entity against entries on various watchlists.

There are many watchlists against which an organization may, for regulatory or risk purposes, be required to screen individuals and entities when initiating a business relationship. These include sanctions lists published by governments or economic, political, and law enforcement bodies. Lists published by commercial sources, such as politically exposed person (PEP) lists, and internal blacklists created by companies themselves may also be required for review.

Sanctions lists contain entries of debarred individuals or entities due to involvement in criminal activity such as money laundering, international terrorism, and financial crime. Sanctions lists may also include lists of embargoed countries.

PEP lists contain entries of high-profile (and often high-value) public figures, such as business leaders, prominent social figures, and members of political parties. Due to their position in society, their political position, or associations and relationships with other high-profile parties, PEPs may be subject to potential bribery or misuse their power and influence for personal gain or financial advantage.

For simplicity, all such watchlists are referred to as *reference data* throughout the remainder of this document.

2.1 High-Level Overview

Oracle Watchlist Screening runs on the Oracle Enterprise Data Quality platform, and therefore shares common hardware. Built upon a modular architecture, it can be tailored to meet the compliance screening processes specific to each industry; the types of records being screened; the match rules applied; and the processes for investigating, managing, and escalating any potential match. Hence Oracle Watchlist Screening includes configurable data optimization, matching, and case management modules to provide tailored solutions to meet the specific needs of each customer.

Data optimization prepares both customer and watchlist data ahead of matching to ensure that any errors and variances in both the structure and content of data are resolved. Matching takes the output from the data optimization module and attempts to determine potential matches between records contained in list datasources and customer data. Those records that exhibit a close match are then presented to case management for investigation and resolution by the compliance team.

An overview of Oracle Watchlist Screening and its functional components is shown in Figure 1.

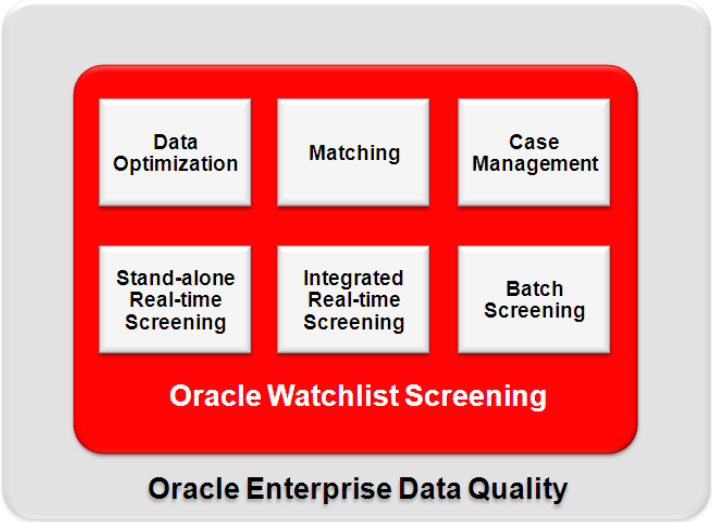

Figure 1. An overview of Oracle Watchlist Screening and its functional components

The underlying Oracle Enterprise Data Quality platform is a state-of-the-art information quality analysis and transformation tool employing a multithreaded, client-server-based architecture built around a Java framework. This allows for simple deployment on a wide range of platforms. The client–server approach enables users of Oracle Watchlist Screening to access multiple instances of the core system while having full flexibility to create new jobs, processes, and reports from their own devices. This architecture also enables the same instance of Oracle Watchlist Screening to be used by compliance teams distributed across multiple locations.

Oracle Watchlist Screening can screen records in both batch and real time, making it suitable for a range of use cases including client on-boarding and regular screening of existing customers.

2.2 Oracle Watchlist Screening Architecture

At the core of Oracle Watchlist Screening is an advanced matching and screening engine, containing a suite of configurable match rules for determining both exact and fuzzy matches between working (customer data) and reference data. A screening client is downloaded from the server and installed on each compliance worker's machine. This enables access to Oracle Watchlist Screening and associated components. Oracle Watchlist Screening also includes an integrated reporting environment and a standard set of reports that can easily be customized and augmented to suit specific customer requirements.

Figure 2 shows each component of Oracle Watchlist Screening in more detail and illustrates how each of these interacts in a standard implementation.

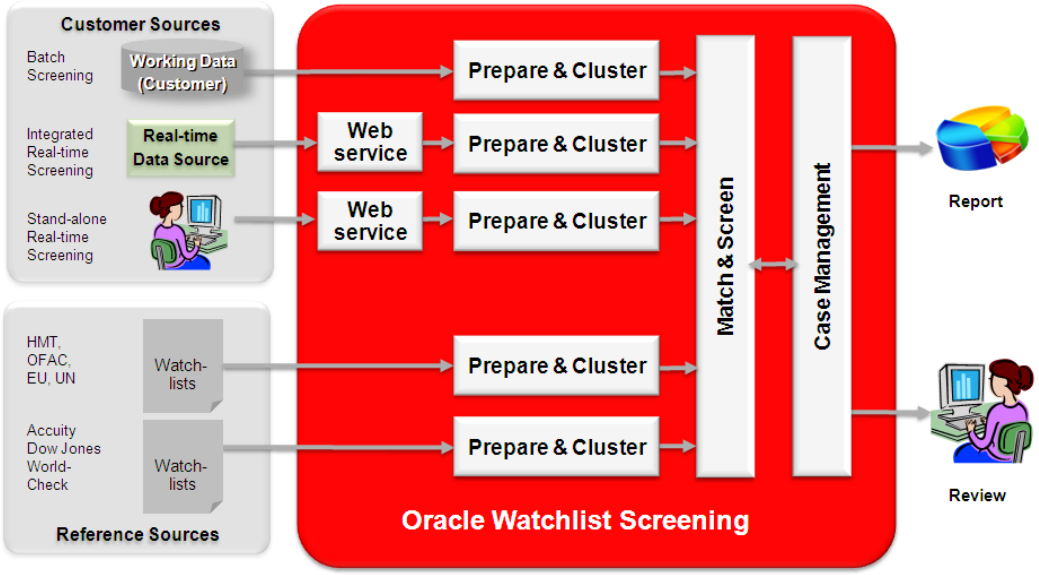

Figure 2. How each component of Oracle Watchlist Screening interacts in a standard implementation

Oracle Enterprise Data Quality is included as part of the deployment exclusively for use with Oracle Watchlist Screening (restricted use license). Other uses of Oracle Enterprise Data Quality require a separate license.

2.2.1 Connecting to Datasources

Oracle Watchlist Screening uses database connectors to connect to both working (customer) data and to each of the external reference datasources, as shown on the left in Figure 2.

Working data may persist in a number of databases and repositories across the organization, often using various file formats. Oracle Watchlist Screening can work with any type of customer data, the preparation of which is normally part of system configuration at the time of implementation.

To make it easier for customers to prepare their data in advance, thereby shortening implementation timeframes, Oracle Watchlist Screening includes a standard customer data interface, specifying a standard file format (.csv) for working (customer) datasources.

Providers of reference datasources publish their list data in various files and formats, and Oracle Watchlist Screening includes a preconfigured connector to each of the most common sources. See section 4.4 for more details.

2.2.2 Preparation and Clustering

Data preparation is a crucial element of the process and has a direct impact on the accuracy of the screening process, the elimination of risk, and the reduction of false positives for the compliance team to review.

For reference datasources, preparation and clustering are carried out as the same process because the list data is presented in a known format. For customer data, preparation and clustering are normally undertaken separately, because the often unstructured nature of this data requires the use of comprehensive on-board utilities to prepare the data ahead of the clustering process.

2.2.3 Matching and Screening

The matching and screening component of Oracle Watchlist Screening provides sophisticated fuzzy logic, combined with advanced data optimization techniques to ensure highly accurate screening. Effective screening relies upon nonexact complex fuzzy matching, because criminal elements targeting organizations often transpose names, dates of birth, and other personal information in an attempt to conceal their true identity.

Oracle Watchlist Screening also uses secondary identifiers as part of the screening process. This allows the software to differentiate between individuals and entities with common names along with a large number of discrete match rules available that can be activated. With the correct application of these to customer and list datasources, false positives can be reduced to a minimum without increasing risk.

2.2.4 Case Management

Potential matches between working and one or more reference datasources are presented as discrete cases within the case management application. Each case is allocated to a workflow, enabling the investigation of each potential match to progress in a structured way. Cases are given a priority using a match strength score (or in some cases, combined match and risk score) and assigned to case reviewers for investigation.

Case reviewers can view each potential match (shown as an alert in Figure 3) along with the match rules that caused the case to be flagged on a single screen. Case management also provides comprehensive audit tools that enable attachments to be appended to capture a full history of all decisions and state changes applied.

An individual case is created where a potential match between a customer record and a list entry is identified. A single match between a customer record and a list entry generates an alert that is progressed and tracked through case management. Cases may contain multiple alerts (for example, where a customer record matches more than one list record).

Cases and alerts are effectively owned by the case management system. The relationship data is owned by the matching process and will be regenerated when a matching process runs. The relationship data is linked with the alert and case data by the alert key and case key.

Multiple names appearing in a single customer or list entry record are wrapped up into the same alert. Note that a case may consist of one or more alerts; however, an alert cannot exist in more than one case because the case key must be the same as, or a subset of, the alert key. This principle is illustrated in Figure 3.

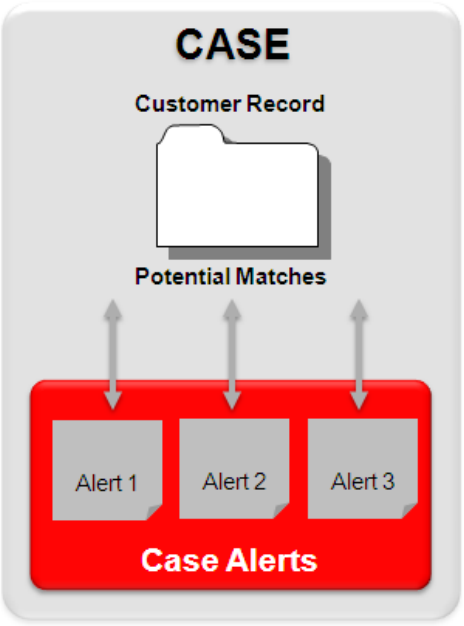

Figure 3. A case may consist of one or more alerts.

2.2.5 Configuration Reporting

Configuration management features provided by Oracle Watchlist Screening make it easy to manage, audit, and share match rules across the organization, ensuring that all business units are fully optimizing their screening processes. The configuration management features enable the user to

- Create "configuration diffs" on an ad hoc basis
- Save the results in a format that can be read using common tools
- More easily diagnose problems in a particular implementation

Screening configurations can be published as reports, making it easy to conduct an audit of match rules applied across a broad spectrum of risk profiles. This provides an extra layer of auditing by making it simple to determine whether and how match rule configurations have changed compared with the configuration at a previous point in time. This is particularly important for demonstrating robust procedures and effective due diligence during any investigation undertaken by the regulator.

2.2.6 Management Reporting

Oracle Watchlist Screening provides comprehensive management reporting to give an enterprisewide view of risk. MI reports can either be created internally using the on-board reporting tools, or the data can be exported to external MI reporting tools via .csv files or as Microsoft Excel spreadsheets. The on-board customizable reporting tools provide a graphical view of the screening results, while also offering access down to an individual case level.

3. Screening Schedules and Data Flows

The screening process should be run on a regular basis in order to detect new potential matches. The frequency of runs will be defined by the compliance team and will be based on their view of risk. The majority of Oracle's customers screen every 24 hours.

Due to the nature of data changing in both the working and reference datasets, it is imperative to screen all customer data against all of the reference data lists, and not just the new and changed records. This is made possible using the powerful matching engine in the Oracle Enterprise Data Quality platform, which is able to process very large datasets on modest hardware. Intelligent algorithms for suppressing repetitive possible matches (often called false positives) reduce the workload on the review user.

Data to be screened can be derived from either real-time sources (for example, where front-office staff are screening walk-in customers), in batch mode during regular screening cycles of existing customers, or as ad hoc screening.

All screening methods can be run concurrently, enabling screening during on-boarding of new customers to be undertaken at the same time that regular checks are being carried out on existing individuals and entities.

3.1 Preparing Data for Screening

Oracle Watchlist Screening employs advanced data-preparation techniques to ensure that customer and list data is optimized to match rules. This is necessary to achieve a high degree of accuracy in the screening process, minimizing high numbers of false positives and the potential for false negatives. Processes to interpret the data, understand it, and optimize it for screening are provided by Oracle Watchlist Screening, including parsing out multiple names in a single field, finding names in address fields, and spotting multiple names in name fields. These are all potential data quality issues.

3.2 Working with International Data

Oracle Watchlist Screening provides optional language packs to enable customers to screen data held in non-Latin script. Language packs provide language-specific rules to enable screening of data in non-Latin format against lists that hold names in Latin form. This normally involves a transliteration processor, with associated reference data, and a dictionary of name variants. The transliteration processor is used in customer data preparation, and the reference data in matching.

Language packs are offered as

- Basic (transliteration only)
- Medium (transliteration and name variant dictionary)
- Complex (third-party data and extensions to handle transcription)

The type of language pack required depends upon the writing system used to capture the original customer record. For example, a basic language pack will be required for screening customer data held

in the Greek alphabet, whereas Arabic would require the complex language pack. Note that Oracle Watchlist Screening supports multiple languages on a single instance of the software.

3.3 Match and Risk Scoring

Match scoring is a function of how close a match exists between an individual or entity and an entry on one or more watchlists—the closer the match, the higher the score. High match score records are flagged as high priority within case management, as those requiring urgent attention. This capability is delivered out-of-the-box as part of Oracle Watchlist Screening.

Oracle Watchlist Screening also offers the option of a risk scoring module to complement its match score capabilities. Risk scoring enables a further layer of data preparation to take place ahead of screening, making it easier to prioritize potential matches in line with the organization's risk tolerance.

Risk scoring works by allocating a value to one or more attributes in both customer data and watchlist entries that may be seen as factors contributing to risk, such as country of residence, occupation, membership of a particular regime, product holdings, and investments.

The compliance user is able to generate a report of match score versus risk score from within case management, with highest priority alerts for review being those with strong match and high risk, as shown in the bottom-right of the grid in Figure 4.

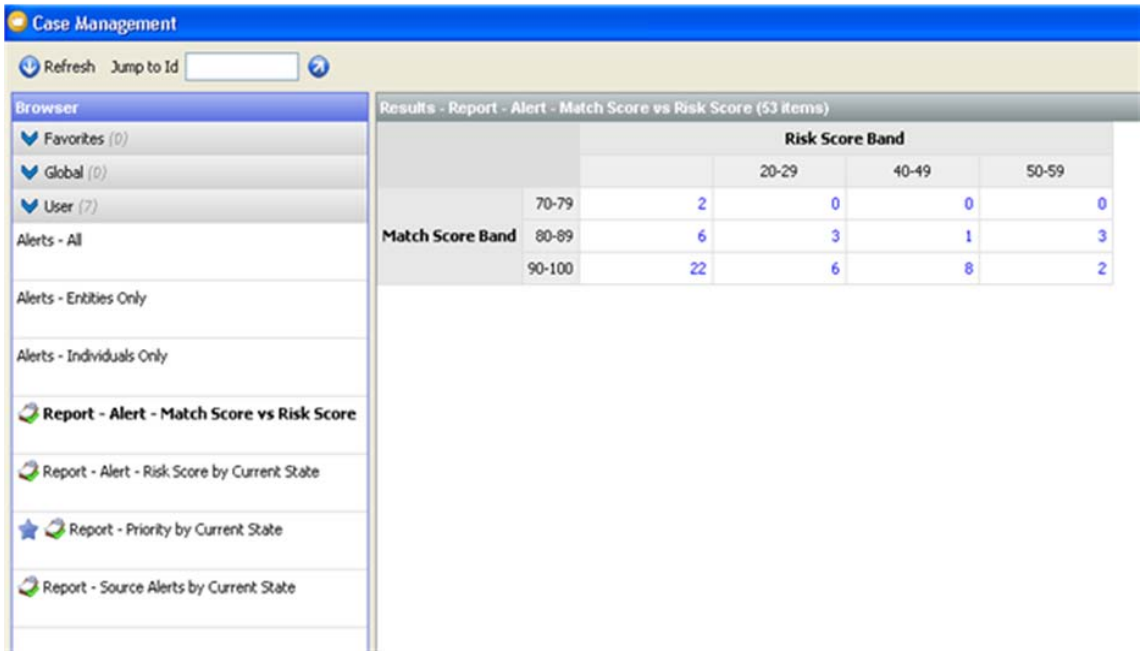

Figure 4. Users can generate a prioritized report of match score versus risk score from within case management.

Should a match be found against any individual or entity, the combined value of the match score and the risk score will be included within the alert presented to case management, and the entry will be prioritized accordingly. This makes it easier to prioritize review activity against those matches that present the greatest potential threat to the organization.

3.4 Batch Screening

Batch screening of working data against reference datasources usually takes place at periodic intervals (for example, on a daily basis, often overnight), with results being made available for compliance teams at the start of the next working day. Both working and reference data are downloaded into a set of repositories via a suite of standard out-of-the-box connectors and are prepared in readiness for matching. The process workflow for batch screening can be summarized as follows:

- The chosen sanctions/PEP lists, referred to as reference data, are read into the Oracle Enterprise Data Quality server for matching.
- Customer (working) data is read into the Oracle Enterprise Data Quality server and prepared for matching.
- The matching process is then run to identify exact and potential matches between working and reference data.
- Matching can be run at the same time as investigators are working on the alert queues.

The compliance team accesses and reviews the results of the matching process using Oracle Enterprise Data Quality client applications—either match review or case management.

3.5 Real-Time Screening

The capability to screen walk-in clients in real time is offered as an application-independent Web service. This gives customers the option of choosing either Oracle's purpose-built, real-time screening UI, or using their own real-time screening client interface and integrating this with Oracle's comprehensive screening engine.

In both cases, the application is accessed via Web service interfaces, which are automatically generated and published, requiring no coding.

3.5.1 Bundled Screening

Oracle's real-time screening UI overcomes the need for customers to create their own UI. This prebuilt interface can be embedded within client applications and offers the advantage of being fully integrated with Oracle Watchlist Screening's case management module.

Bundled screening exhibits the following characteristics:

- Customers can include additional characteristics to allow risk scoring to occur at the point of screening.
- Match rules and case workflow configurations can be created to screen a range of risk profiles, regardless of whether Oracle Watchlist Screening is running on a single server or instances of multiple servers.

- Users with appropriate permissions can gain access to multiple projects, with access to projects being governed by the user's project rights to protect customer confidentiality.
- Oracle Watchlist Screening can be configured to allow the UI to capture additional customer data to be passed to the system for review and audit purposes.

3.5.2 Custom Integrations

Oracle Watchlist Screening also exposes Web services interfaces to support customers who wish to continue to use their own bespoke UI for real-time screening. This enables screening processes to be tailored to suit business practice, such as how to process records when a potential match is found, and the level of information disclosed to the user.

- Oracle Watchlist Screening can easily be integrated and used within an existing front-office application, new dedicated application, or portal Web page and is accessed.
- An application is included in order to test and demonstrate the real-time screening service.
- In real-time mode, records are passed to Oracle Enterprise Data Quality via a Web service interface or by using the Java Message Service (JMS).

3.5.3 Real-Time Screening in Action

Figure 5 illustrates a real-time screening workflow where an organization chooses to integrate Oracle Watchlist Screening with its own customer on-boarding system.

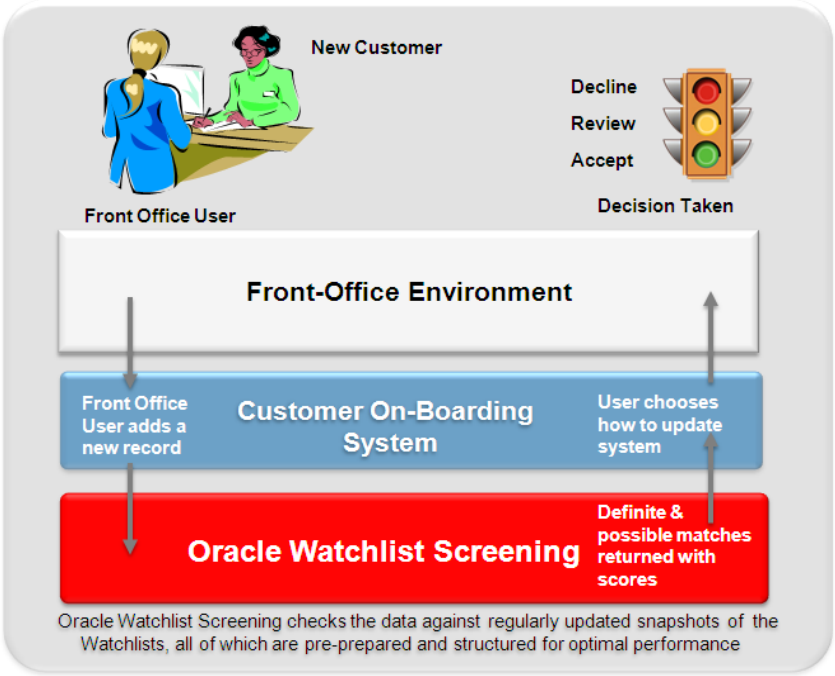

Figure 5. Oracle Watchlist Screening is integrated with a custom on-boarding system.

3.6 Ad Hoc Screening

Ad hoc screening is often carried out at the end of each day, when lists of individuals and entities are submitted to a central team of compliance workers for screening. These records are then uploaded for screening as a batch file. Real-time screening refers to the screening of walk-in customers at the time they make a request for service, such as for parcel shipment or account opening. For all intents and purposes, ad hoc screening follows the same process as for batch screening, and is possibly best thought of as batch screening on demand.

4. Working Data

4.1 Supported Data Formats

Because Oracle Enterprise Data Quality is a data management platform, it has been designed to read data in an extremely wide variety of formats, including text files (delimited, fixed width, XML), document formats (Microsoft Excel), and RDBMS (using JDBC connectivity).

Put simply, the required format is "whichever format is easiest or best suits your organization." Internal mapping and the manipulation of fields allow a very flexible interface definition, enabling field naming and content to be easily resolved. Customer data formats are converted to those used in the matching process as part of the implementation of Oracle Watchlist Screening. During this transformation process, data quality issues—such as superfluous data after surnames—are assessed and rectified.

The Oracle software will not change the source data during its processing; all changes occur internally to Oracle Watchlist Screening. Importantly, Oracle Watchlist Screening takes the raw data in the format most convenient for you. There is no requirement for a costly or highly defined data feed. The data used for matching within Oracle Watchlist Screening is categorized as either an entity or an individual, defined as follows.

4.2 Entities

This category covers entities including organizations, companies, charities, and political parties.

Table 1 shows the attributes that are mandatory (M) and optional (O) when screening records of entities against watchlist sources. Where risk scoring is deployed, additional data attributes (X) can be used in the matching process to assist in prioritizing review work in line with criteria defined by the compliance team. Underlying directors or beneficial owners are treated as individuals for the purpose of screening.

DATA FIELD	TYPE	RISK SCORING
Unique ID	M	
Company Name	M	
City of Operation(s)	O	
Country of Operation(s)	O	X
Full Operating Address	O	
Country of Incorporation/Registration	O	X
Company Registration Number	O	
Industry Code/Type	O	X
Product(s)	O	X

Table 1. Certain criteria are mandatory (M) and others optional (O) when screening entities for risk.

4.3 Individuals

Similar to the screening for entities illustrated in Table 1 above, Table 2 shows the attributes that are mandatory (M) and optional (O) when screening records of individuals against watchlist sources:

DATA FIELD	TYPE	RISK SCORING
Unique ID	M	
Title	M	
Given Name(s)	M	
Family Name	M	
Full Name	M	
Date of Birth	O	
Gender	O	
City of Residence	O	
Country of Residence	O	X
Full Residential Address	O	
City of Birth	O	
Country of Birth	O	X
Nationality	O	
Passport Number	O	
National Identity Number	O	
Product(s)	O	X
Occupation	O	X

Table 2. Certain criteria are mandatory (M) and others optional (O) when screening individuals for risk.

Additional information specified by your compliance team—such as "date of registration"—can be added to the working data and will be available for reporting review purposes. Supplying the data to

the system is usually simply a matter of placing a file in an agreed network location or defining a SQL query for the system to execute at the scheduled time.

4.4 Interface Techniques

The Oracle Enterprise Data Quality platform supports both "push" and "pull" interface methods. The majority of customers adopt a push approach using a file-based mechanism to transfer data to Oracle Watchlist Screening. Oracle supports many file formats, including XML, Delimited Text (.csv, tab, and so on) and Microsoft Office. Very few customers allow the software to pull data from their source system, and where this does occur, the software normally connects to a dedicated table within a warehouse using JDBC to retrieve the data.

4.5 Reference Data

Oracle Watchlist Screening works with all national, regional, and global sanctions and watchlists, including but not limited to

- Government lists
  - Her Majesty's Treasury (HMT) sanctions list
  - U.S. Office of Foreign Asset Control (OFAC) sanctions lists
  - European Union (EU) sanctions lists
  - UN (United Nations) sanctions lists
- Commercial watchlists including those provided by
  - Accuity
  - Dow Jones
  - World-Check
- External Risk Score Providers
  - Safe Banking Systems (SBS)

Customers' own blacklists or other denial lists can also be included and used as reference data for screening through simple configuration changes.

For all major open source and commercial sources, Oracle has designed preconfigured connectors that are able to read in the data from the source and prepare it for matching. It is also possible (if required) to automate the collection of this watchlist data. A process is loaded within Oracle Watchlist Screening to enable the connector to reach out and read the list data.

While Oracle provides the required connectors as part of Oracle Watchlist Screening, it is the customer's responsibility to sign a contract with each of the commercial list providers in order to gain access to the list data. Open source lists (such as those provided by government agencies) are available as an anonymous download and require no subscription or contract.

5. Hardware and Software Requirements

Oracle Watchlist Screening can run in both Windows and non-Windows environments. Figure 6 shows typical environments for hosting each of the solution components, along with the protocols used for message communication.

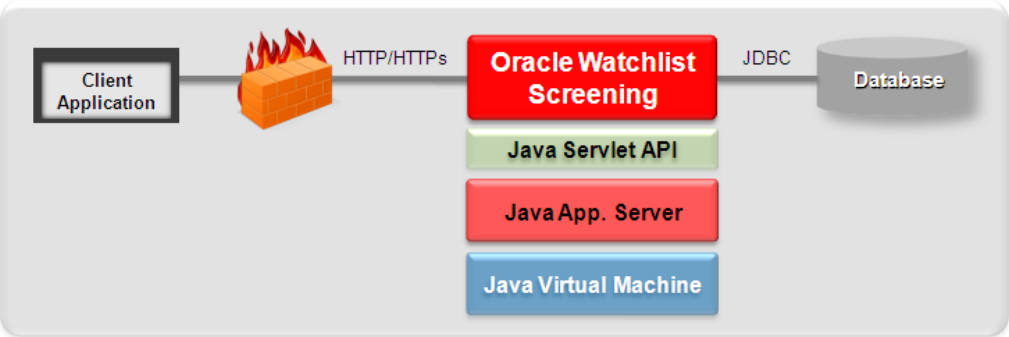

Figure 6. This figure illustrates typical environments for hosting each of the solution components, along with the protocols used for message communication.

Table 3 gives details of the hosting environments and operating systems (OSs) for each domain. The following sections provide more details of the specific OSs and versions supported.

DOMAIN	HOSTING ENVIRONMENT	OPERATING SYSTEM
Client Application	• Java Web Start	• Windows • Linux
Java Application Server	• Tomcat • Oracle WebLogic • IBM WebSphere	• Windows • Linux • AIX • Solaris
Oracle Watchlist Screening	• Java	• Windows • Linux • AIX • Solaris
Oracle Enterprise Data Quality Repository	• Oracle • PostgreSQL	• Windows • Linux • AIX • Solaris

Table 3. Each domain has specific hosting and OS requirements.

5.1 Windows Environment

Oracle Enterprise Data Quality is a Java Web application using a Java Servlet engine and a Java Web Start GUI. It has been certified to work on Java 1.5 and 1.6 and runs on the Windows platforms listed below:

- Windows Server 2003
- Windows Server 2008
- Windows 7

5.2 Non-Windows Environment

It is expected that any compliant Java Servlet 2.4 engine will successfully support Oracle Enterprise Data Quality, as will the Postgres and Oracle RDBMS systems. The following platforms have been certified as supporting Oracle Enterprise Data Quality:

- Tomcat
- Oracle WebLogic 10 (or later)
- IBM WebSphere 7.0

6. The Oracle Enterprise Data Quality Repository

The Oracle Enterprise Data Quality repository is a fundamental part of Oracle Enterprise Data Quality and serves a number of purposes, including

- User administration
- Temporary storage of working and reference data
- Management of decision persistence
- Management of the decision audit trail

For security and functionality reasons, Oracle Enterprise Data Quality uses an RDBMS (Oracle and Postgres are currently supported) for its data repositories and uses Postgres as the default repository.

The repository is a component of Oracle Enterprise Data Quality and is created and initialized during installation. There is no requirement for dedicated maintenance/intervention from a DBA, and no specialist knowledge is required to support the repository because it is automatically managed and maintained by Oracle Enterprise Data Quality.

7. Disaster Recovery and High Availability

Every company has its own disaster recovery strategy with specific requirements. However, in the vast majority of implementations, the following information holds true:

- The business process does not as a general rule require a high level of availability. For example, if the server fails and a new machine cannot be configured for 24 hours, this is usually acceptable to the compliance team.
- In the case of failure, a new application server can be installed and pointed at the Oracle Enterprise Data Quality repository to restore service.
- It is acceptable to have a warm DR server with the Oracle Enterprise Data Quality server installed and configured to allow a rapid restoration of service in the event of system failure.

The majority of system data does not require specific backup, because it can be re-created (by taking another extract from the customer database or redownloading reference datasources). The exception to this is the Oracle Enterprise Data Quality repository. This should be part of a standard backup regime because it contains information that is only maintained within Oracle Watchlist Screening—for example, user administration, match decisions, and audit trails.

Organizations will normally have their own tools and processes for database backup. Otherwise, this can be performed using a simple file-copy mechanism.

Appendix 1: Glossary of Terms

TERM	DEFINITION
Ad hoc screening	Screening that takes place outside of normal processes.
Alert	A potential match between a customer and watchlist entry presented to case management for review.
Batch screening	Periodic screening of a customer (or working data) against reference datasources.
Blacklists	Customer's records of individuals and entities denied a service.
Fuzzy matching	Matching based on reasoning that is approximate rather than accurate.
Case	A collection of related alerts with a distinct lifecycle. A single case may be generated where a customer record matches one or more watch-list records.
Case management	A suite of integrated tools for investigating and reporting on potential matches.
Configuration comparisons	Used to highlight the similarities and differences of configuration between two components or sets of components.
.csv	Comma Separated Values—a delimited data format that has fields/columns separated by the comma character and records/rows terminated by new lines.
Customer data	Records of individuals and entities held by the organization.
Data optimization	The processes applied in preparing data in readiness for screening (for example, profiling, parsing, and deduplication).
False positive	A potential match that is not true but is presented as such.
False negative	A true match that was not identified.
Matching	A comparison of two or more records to determine whether or not a relationship exists.
Match score	A numeric value attributed to a particular match rule that caused two records to be related.
Match review	A process within Oracle Watchlist Screening for investigating potential matches.
MI reporting	Reports presented in a style and format suitable for executive management.
Out-of-the-box	Preconfigured and ready to install.
Politically exposed person (PEP)	A current or former official of government, a major political party, or certain corporations and his/her immediate family members, as well as any close personal or professional associate of such a person.
Raw data	Data that has not been optimized or prepared for screening.
Real time	In the context of screening, checking for potential matches between customer and reference data on request.
Reference data	Watchlist data that is used in matching and is to be screened against.
Risk score	A score calculated for each individual or entity on each watchlist, based on various attributes such as country of residence, operating country, associated regime, and so on.
Sanctions lists	Lists published by government agencies (for example, HMT and OFAC) containing records of sanctioned individuals and entities with whom organizations are prohibited from conducting business.
Secondary identifiers	For example, date of birth, name in original script, country of residence, and photographs.

Screening	The process of determining whether a potential match exists between a customer record and one or more entries contained within list datasources published by government agencies and commercial providers.
Screening client	A client application executing on the user's machine.
Watchlist	A list available from a provider containing details of any/all of the following: sanctioned individuals, entities, PEPs, and their relatives and close associates.
Walk-in customers	Customers not having a preexisting account when requesting a service or opening an account.
Working data	Customer data held by the organization that is used in matching and is to be screened.

Appendix 2: Installing Oracle Watchlist Screening

A version of Oracle Enterprise Data Quality needs to be installed on the server machine prior to loading Oracle Watchlist Screening. Oracle Enterprise Data Quality can be deployed into both Windows and non-Windows environments.

The installation process for Oracle Enterprise Data Quality has been designed to minimize impact on IT resources. The Oracle Enterprise Data Quality software consists of a server component and a client application. The installation will install both components onto the target machine, which will become an Oracle Enterprise Data Quality server. The client machine runs the Java Web Start process and requires Java 5 or Java 6 to be installed. See the following documents for detailed instructions on how to install Oracle Enterprise Data Quality in both Windows and non-Windows environments:

- "Oracle EDQ Installation Notes"
- "Oracle EDQ Advanced Installation Notes"

Oracle Watchlist Screening can be loaded once a version of Oracle Enterprise Data Quality has been installed. It uses custom widgets (preconfigured processors), gadgets (match extensions), and database connectors to extend the functionality of Oracle Enterprise Data Quality. These components are provided as Java Archive (JAR) files, which can be installed using the Oracle Enterprise Data Quality extensions configuration tool, accessed via the server configuration area of the Oracle Enterprise Data Quality launchpad. Also, a set of online help files and a customized version of the case management extended attributes file (flags.xml) are provided as part of Oracle Watchlist Screening.

See the "Oracle Watchlist Screening Implementation Guide" for detailed instructions.

Appendix 3: Example Implementations

The following example implementations are based upon internal benchmarking. Note that variable factors such as desired processing window, screening frequency, and the rule set deployed influence actual performance. Typical deployment options are shown in Figure 7.

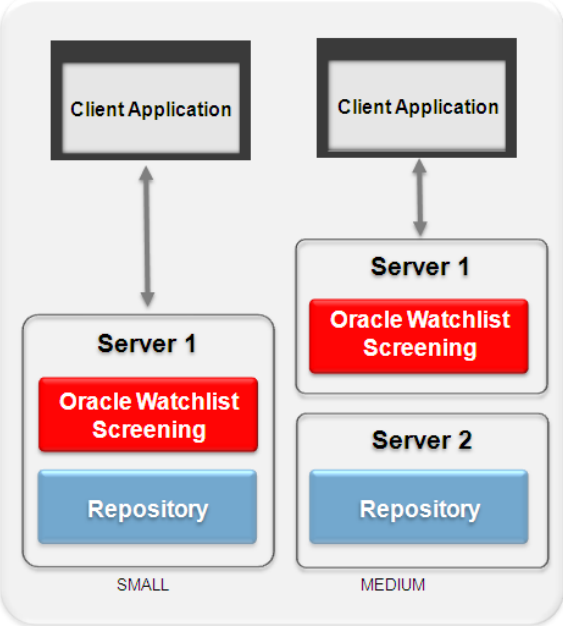

Figure 7. Typical deployment options for Oracle Watchlist Screening

Typically, small batch screening deployments (up to 5 million records screened) can run on a single server, with both Oracle Watchlist Screening and the Oracle Enterprise Data Quality repository hosted on the same system. Medium deployments (up to 10 million records screened) will run on dual servers, one hosting Oracle Watchlist Screening and the other hosting the Oracle Enterprise Data Quality repository.

Beyond this, there are a number of deployment options, including taking into account the volume of records screened and the requirements for hardware redundancy. These options are best discussed with Oracle's sales consultants. In addition, as volumes increase, or as the available time period for processing decreases, the following scaling options can be considered:

- Increase the number of processors and processing cores
- Increase server memory (RAM)
- Split the application server and the repository server onto separate servers

In general, for most installations we recommend the "medium" specification server configuration available at time of purchase. This will usually provide scalability and allow for expansion.

Typical Batch Screening Configurations

Typical server configurations are given for both the small and medium deployments referred to in Figure 7.

UP TO 5 MILLION RECORDS, 3 CONCURRENT USERS		
Processor Cores	RAM	Storage
4	4	250 GB

In this deployment, a single server would host the Web server, Oracle Watchlist Screening, and the repository server.

UP TO 10 MILLION RECORDS, 7 CONCURRENT USERS		
Processor Cores	RAM	Storage
8	8	500 GB

This deployment would be split across two servers. One server would host both the Web server and Oracle Watchlist Screening; the other would host the repository server.

Typical Real-Time Screening Configuration

Real-time screening can also be referred to as *single name check*. The table below specifies the minimum recommended configuration for real-time screening:

MINIMUM SERVER CONFIGURATION		
Processor Cores	RAM	Storage
4	4	250 GB

In this deployment, a single server would host the Web server, Oracle Watchlist Screening, and the repository server.

The actual system required will be a function of the number of concurrent real-time users, the numbers of records to be screened, and the response time required. To ensure that performance expectations are met, Oracle recommends a discussion with our sales consultants, who will be happy to assist with system dimensioning ahead of any deployment.

Deploying into a Virtual Environment

When planned and deployed properly, virtual machines (VMs) that have been consolidated onto a physical server should operate normally and provide a level of service that is indistinguishable from their physical counterparts. Unfortunately, this doesn't always happen when deployed applications have substantial resource requirements unless the VM has been configured to expect this type of resource utilization.

Oracle Enterprise Data Quality has been designed to fully exploit the physical resources it has been allocated. It is not unusual, therefore, for an Oracle Enterprise Data Quality batch job to consume the maximum CPU resources available for hours at a time. This might precipitate virtual server

performance issues if the VM has not been configured to allow for this. In this case, the VM may throttle back the performance of Oracle Enterprise Data Quality by applying load balancing rules that do not recognize the characteristics of a typical Oracle Enterprise Data Quality job.

When deploying Oracle Enterprise Data Quality in a VM environment, administrators must ensure that the VM is configured to permit Oracle Enterprise Data Quality to have at least the minimum recommended resources dedicated to it. Configuring the Oracle Enterprise Data Quality repository on a virtual disk array should also be avoided, to prevent throttling of disk I/O and reducing the performance of Oracle Enterprise Data Quality jobs. The Oracle Enterprise Data Quality repository may be configured on a network-attached server, but it should be a real rather than a virtual array.

Oracle Watchlist Screening:
Technology Guide
August 2011

Oracle Corporation
World Headquarters
500 Oracle Parkway
Redwood Shores, CA 94065
U.S.A.

Worldwide Inquiries:
Phone: +1.650.506.7000
Fax: +1.650.506.7200

oracle.com

Oracle is committed to developing practices and products that help protect the environment

Hardware and Software, Engineered to Work Together