evidian

ID Synchronization for a multi-directory identity repository

Bull
atos technologies

Introduction

Building an identity repository is at the heart of identity and access management. In fact, no matter the quality of an access security policy, no matter the effectiveness of its implementation, and no matter how refined the reporting tools are, if the user-defining data is not reliable, the entire structure is bound to collapse.

At the heart of identity and access management

Building an identity repository is at the heart of identity and access management. In fact, no matter the quality of an access security policy, no matter the effectiveness of its implementation, and no matter how refined the reporting tools are, if the user-defining data is not reliable, the entire structure is bound to collapse.

A number of obstacles may be encountered while building such a repository, like defining a too-complex model, adopting a purely technical approach or even mixing user-defining data and rights-defining data. Before embarking on such a project, it is important to use an approach which will enable you to define simple, robust and effective models and processes during its implementation.

Why a coherent repository is essential

Evidian recommends the use of a model which allows you to separate identity information from access-rights-defining information and the procedure used to build them.

Figure 1. From authoritative data to assigned data

Identity-defining information is referred to as "authoritative data", while access-rights-defining information is part of "assigned data".

	Definition	Examples
Authoritative data	Data which defines identity in the broad sense of the user.	Surname, first name, company identifier, organization, job, position in the company, site, office, phone number, etc.
Assigned data	Data which defines user access rights on target systems and applications.	Application logins and passwords, privileges, timetable for use, LDAP group membership, etc.

Table 1. Authoritative and assigned data

A user's assigned data is generally obtained by applying security policy rules to the authoritative data of the same user. For example, a sales engineer (authoritative data) will have a login and a password (assigned data) to access the application meant for salespersons (security policy).

Located at the beginning of the process, authoritative data is, therefore, at the heart of identity and access management procedures. Any creation, modification or deletion of authoritative data may have an immediate impact on assigned data.

Building an identity repository

There are several ways of building an identity repository for an organization.

For example, the already existing identity data management system may be simple enough for an identity management workflow to supply data to a single directory shared by all the organizations (more information on our website www.evidian.com).

But in most cases, the organizations have deployed several directories, which collect identity data in formats that are often incompatible and difficult to share. Moreover, they are managed by groups of dedicated administrators, who apply their own data model.

For example, within the human resources management systems, administrators are tasked with managing users' names, jobs and the attributes relating to their organizations; in the PBX database, administrators are in charge of managing telephone numbers; in LDAP directories, administrators must manage users' e-mail addresses, etc.

This incoherent situation results in high non-quality costs:

- A loss of productivity for the IT teams, which must manage the same data in several places and in several formats
- Uncontrolled and attack-prone identities.
- Proliferation of unreliable identities which complicates authorization management and control.

To solve these problems, Evidian has introduced an identity synchronization function (ID Synchronization) that makes it possible to build a source of identity in a single repository.

Identity synchronization

Identity synchronization allows you to create a reliable and coherent **identity repository.**

The ID Synchronization module that offers this feature interacts with other functions of Evidian's identity and access management solution **(Identity Governance and Administration):**

The policy manager uses the repository thus created to initialize access rights creation operations based on the security policy.

The workflow completes the consolidation mechanisms by implementing validation procedures for the most critical applications.

"Provisioning" uses identity data to create, modify or delete access rights in target systems and applications.

Enterprise single sign-on (SSO) uses the data to validate a user's identity.

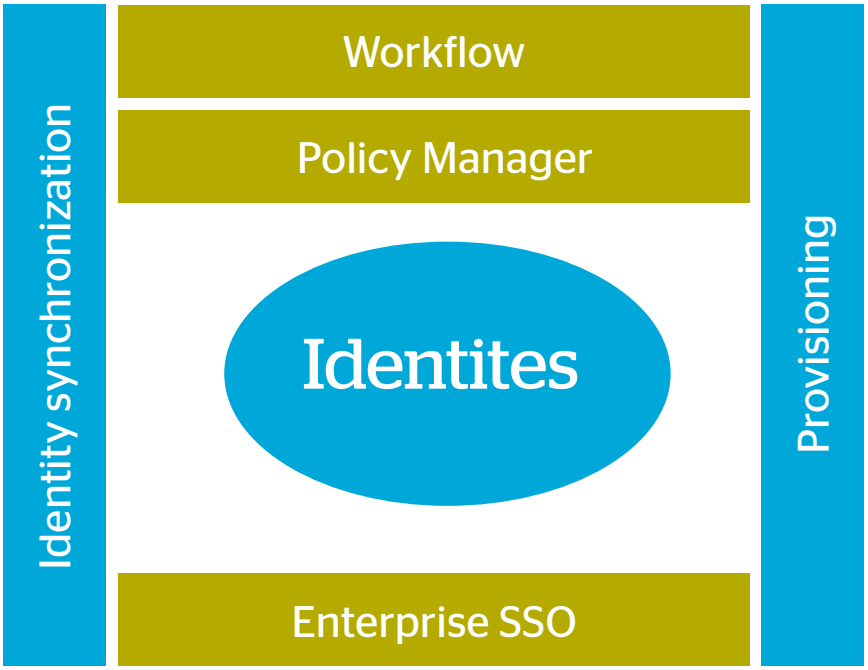


Figure 2. Identity synchronization

Main functions of identity synchronization

Identity synchronization enforces synchronization policies, coupled with the interface of identity and access management modules.

Synchronization functions

Defining and enforcing a multi-repository synchronization policy.

Activating downstream functions

Activating rights update operations through "professional" rule mechanisms within the policy manager. These operations may lead to the activation of provisioning processes on the target applications in the information system.

Logging

Logging identity management operations for analysis and reporting.

Synchronization policy

Generally, the synchronization policy's application range must be restricted to user-defining authoritative data.

This policy will allow the application of the following rules:

- Rules for **consolidating** an identity from several recordings; identity-related data may be located in different repositories
- Rules for **reconciliation** in case of data inconsistency
- Rules for **creating or deleting** a user's data
- Rules for **managing the thresholds** that prevent alteration of the identity repository, in case the data sources are corrupted.

Synchronization may take place in batches or continuously; in any case, it is only based on the latest modifications.

If the number of operations to be carried out exceeds the threshold, the processing is not realized and an administrator is immediately notified. He can replay the stream in simulation mode to analyze the origin of the anomaly.

Corrupted sources can then be corrected and the processing can be restarted.

For the existing solution to be permanently maintainable, it is very important to separate authoritative identity data from assigned data, and to apply synchronization to authoritative data only. For instance, although the reconciliation mechanisms are the same for both data families, they are not subject to the same rules:

- The reconciliation rules for authoritative data depend on the level of trust or reliability associated with each set of data. If a user's telephone number is different in two directories, you only need to define a rule for determining the reference data.
- The reconciliation rules for assigned data (data generally provisioned on target systems) compare the reality of data on a target system with the value that it should have according to the access security policy. They must, therefore, interface with the policy engine. The result of a reconciliation in favour of the actual data may be a change in the access security policy.

Main functions of identity synchronization

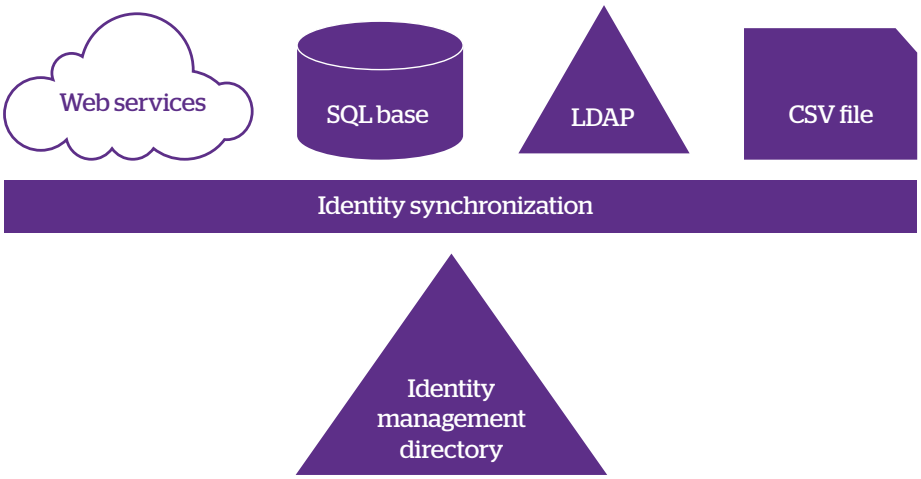


Figure 3. The repositories to be taken into account

Repositories

Identity synchronization works on different types of data sources.

The most common technologies are LDAP directories, relational databases, flat files (CSV, LDIF etc.) or even through Web services.

In a complementary way, it is possible to find interfaces to the databases of Human Resources applications. These HR databases may contain necessary information for defining users or can initiate events for creating or deleting a user's identity within the information system.

If the global identity and access management solution includes its own identity base, identity synchronization must naturally integrate this base into its synchronization mechanisms.

A special case: target application and system repositories

In some cases, the internal repository of the application itself may contain some identity data. Identity synchronization may then use the technical mechanisms, such as agents and connectors, generally used by downstream provisioning to integrate the target data into the identity repository.

In fact, these connectors and agents use public and regular interfaces (API) supplied by the application provider. Instead of giving direct access to the application's internal repository, the use of these public APIs enhances the stability of the installed solution.

A multi-level organization

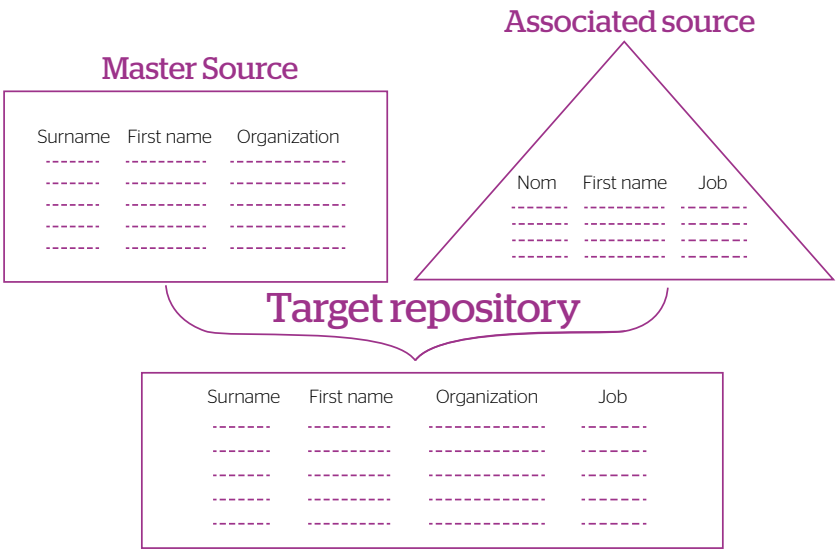


Figure 4. Join operations

To create consistent sets of source data, it is possible to associate a set of "associated" source repositories with a "master" source repository (through joint mechanisms). A "master" source contains the record to which the associated source data may be assigned. Deleting a "master" record will therefore completely erase the record, whereas deleting an "associated" record will be considered as absence of data to be processed, if necessary, using the synchronization rules.

This first aggregation level is used to organize the data so as to have the same view of data from different sources. In this case, the synchronization rules may be top-level rules and may be applied to all the accessible data.

Updating target repositories

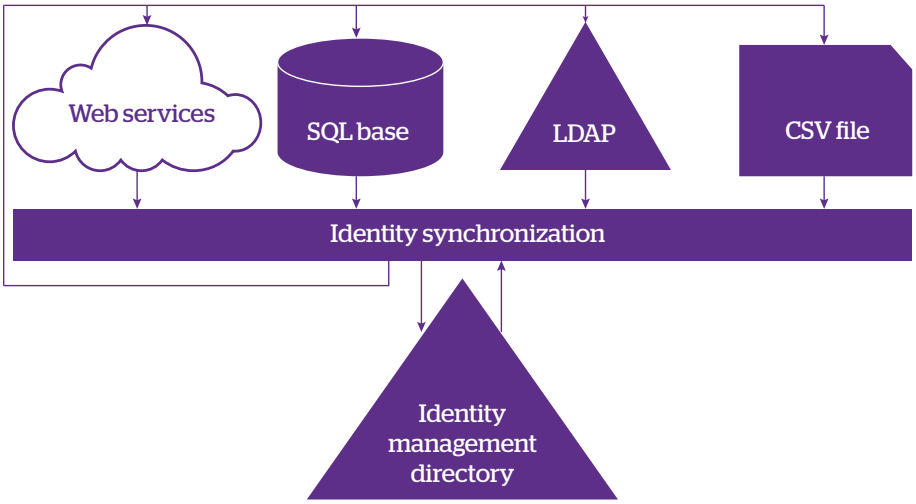


Figure 5. Target repositories

You can use synchronization rules to create a single, centralized repository, which is actually the main objective of the operation.

You can also use them to update a group of target repositories and thus create a coherent identity on a distributed set of repositories.

This distribution of identity enables you to simply solve problems of directory architecture, network flow optimization or even data homogenization in different and/or incompatible repositories.

Rules typology

Quite paradoxically, you can implement identity synchronization just using four basic rules:

Join

This rule allows you to base the recording of an "associated" source on one or more "master" sources.

Attribute Mapping

You can use this rule to define a correspondence relation between attributes from different sources. This relation is well-ordered and leads to an attribute update. This update can also result from the application of an intermediate transformation function.

Finally, this correspondence relation enables you to update multi-value attributes from several sources. A simple example of multi-value attribute is the e-mail address attribute:

In fact, an employee can have several e-mail addresses within an organization: firstname.surname@organization.com, f.name@organization.com, fname@organization.com, firstname.com@country.organization.com, First name/ Name/ Org/Country.... Each e-mail address is available in a specific directory. They can be consolidated within a single e-mail address repository through a multi-value attribute.

Creation

If a user exists in a "master" repository but not in an "associated" repository, you can use this rule to create all the attributes associated with this user, in the "associated" repository.

Deletion

This rule automatically deletes a user's attributes from an "associated" repository if this user has been deleted from the "master" repository.

Other rules

You can work out other rules for creating attributes in repositories. These other rules generally concern the creation of assigned data. Therefore, they are naturally integrated into the policy manager, from which you can create and provision the user access rights on the target applications and systems, using the identity data and in keeping with the policy.

Examples of technical architectures

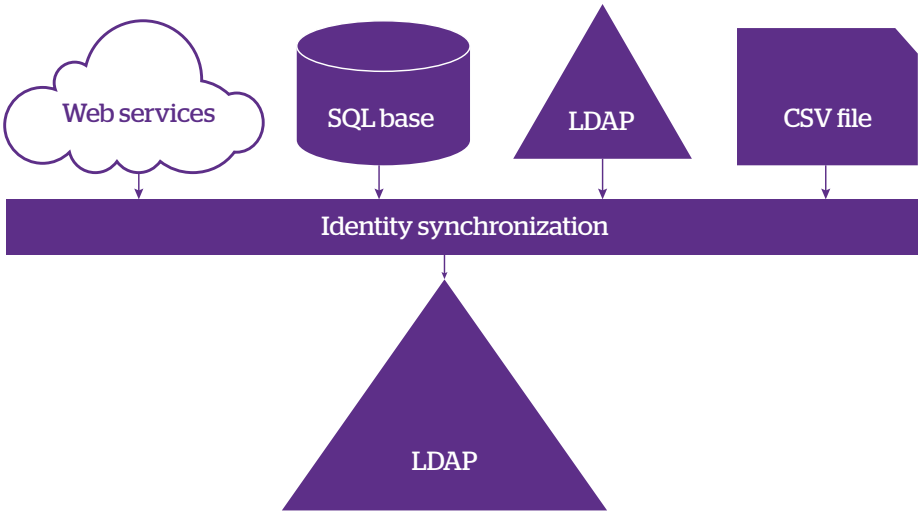


Figure 6. Creating a reference LDAP directory

Case 1: creating a reference LDAP directory

In this case, you can create and consistently maintain the reference identity LDAP directory from different types of information sources (SQL database, CSV files, other LDAP sources, etc.).

These information sources are generally under the responsibility of an HR manager, a partnership manager, or are part of an ERP. The data inside the reference directory result from the aggregation of source data consolidated during join operations.

In this way, you can create a repository which will enable you, for instance, to define an LDAP directory for Enterprise SSO.

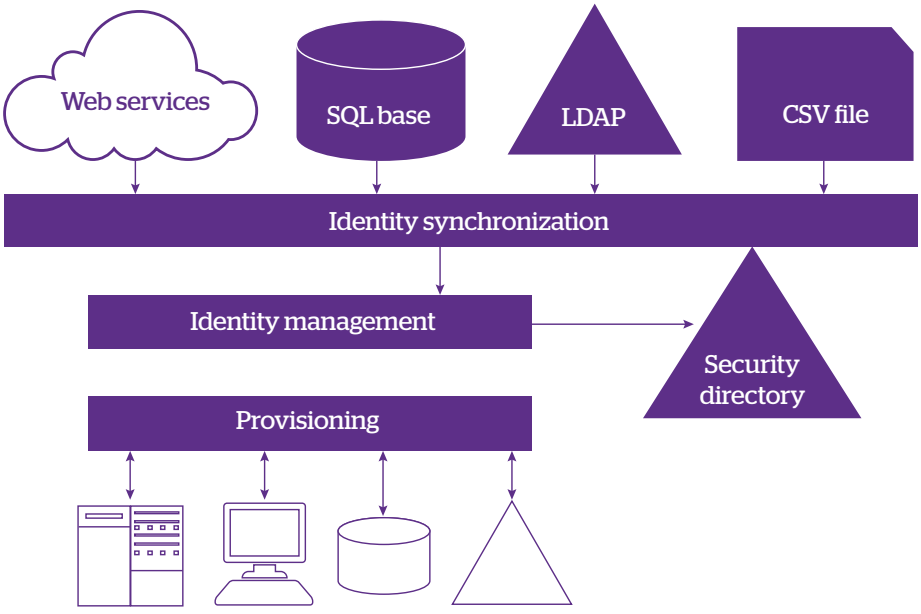


Figure 7. Using an identity and access management solution's database as a single identity repository

Case 2: creating an identity repository using the database of an identity and access management solution

In this case, identity synchronization allows you to fill and consistently maintain the identity base with an identity and access management tool. In general, this base is managed by an identity management module and is used by a provisioning module.

This base may be called a security repository. It can contain all user data or only a subset of user data (see case 3 below).

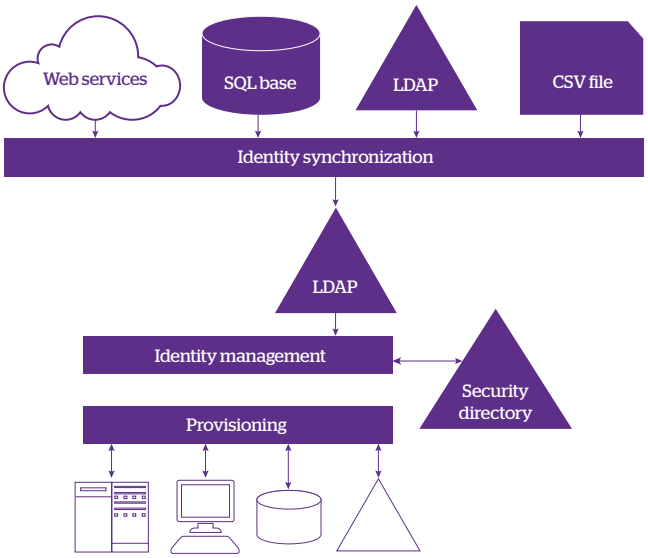


Figure 8. Using the LDAP directory as single identity source

Case 3: using the LDAP directory as identity repository

This is an extension of the previous case. This is the case where the identity and access management solution is using an external LDAP directory as identity repository.

The security repository may then contain, for instance, the access rights data or even additional identity data.

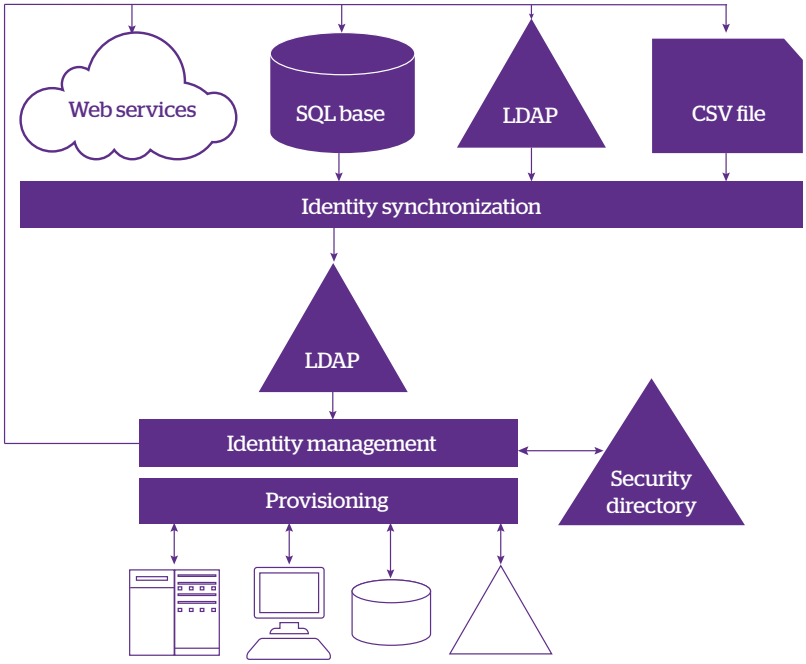


Figure 9. Synchronizing with target-application data

Case 4: synchronizing with target application data

This is the most comprehensive case for which some user data in the target applications' internal repositories may be used as authoritative data sources.

It is then possible to use agents and connectors to synchronize the target applications' data with the other different authoritative sources.

An example of implementation

The following example is a real use case of implementation with Evidian's IAM solution. It uses identity synchronization as well LDAP / security-repository coupling. Based on the technical characteristics of identity synchronization, it is possible to describe the following case of use:

Users are defined in 4 source databases:

- The mainframe repository, which contains an employee's surname, first name and login
- The HR repository, which contains additional data such as e-mail address, location, telephone number, and fax number
- The exceptions CSV file, which contain the list of a user's e-mail addresses. This list contains the log of a user's e-mail addresses following various acquisitions, consolidations and domain changes.
- A web-based database containing the professional mobile number.

The main target database is Active Directory, which contains the list of local Active Directory groups. The master repository is the mainframe repository. The presence or absence of a user in this repository determines the presence or absence of said user in Active Directory.

Data update rules are based on four types of flows.

Creating users

Users are created in Active Directory with all their attributes. The first name, surname and login come from the mainframe repository, and additional data such as e-mail addresses, the location or telephone number come from the HR database. The professional number comes from a database accessible via web services. If a user has several e-mail addresses, the multi-value field "e-mail address" in Active Directory

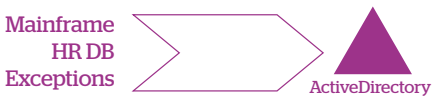


Figure 10. Active Directory content is created from data in the source databases

Creating local groups

The organization's groups, which exist in the mainframe repository, are created in Active Directory. These groups are used to define a set of associated application resources.

Creating local Active Directory groups in the mainframe repository

To associate a global group with a local Active Directory group, the mainframe administrator needs to retrieve the local Active Directory groups on the mainframe. These Active Directory groups must, therefore, be copied to the mainframe.

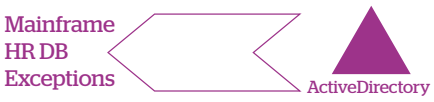


Figure 11. Synchronizing with Active Directory data

Creating group membership relations

Relations of users' membership of local Active Directory groups and global groups, as well as of global groups' membership of local Active Directory groups, are defined in the mainframe repository by administrators. These relations are copied to Active Directory to allow the control of accesses to the corporate resources at the Windows level.

Defining and applying a group-based security policy

Although the most comprehensive user definition is located in Active Directory, the management of global groups is defined in the mainframe. Membership relations are defined by the mainframe administrator. This policy is then "copied" to Active Directory.

The mainframe repository contains a subset of attributes used to perform this group management.

Applications which use Active Directory

Some infrastructure applications may natively use the data available in Active Directory.

Active Directory is, thus, used to manage access to shared disk and printer resources. Moreover, the White Page application uses the data defined in this directory to publish users' public information.

An Enterprise SSO solution can also use this user database to control access to target applications.

No change of directory architecture

In this case of use, the organization did not wish to change the Active Directory schema. The security repository of the identity and access management solution was used to extend the architecture outside Active Directory (cf. case 3 of technical examples).

A solution which adapts to organizational constraints

An identity synchronization solution adapts to technical and organizational constraints.

For example, a group-based security policy can be applied to any repository in the company. It only needs to contain the authoritative information required to define and apply this policy.

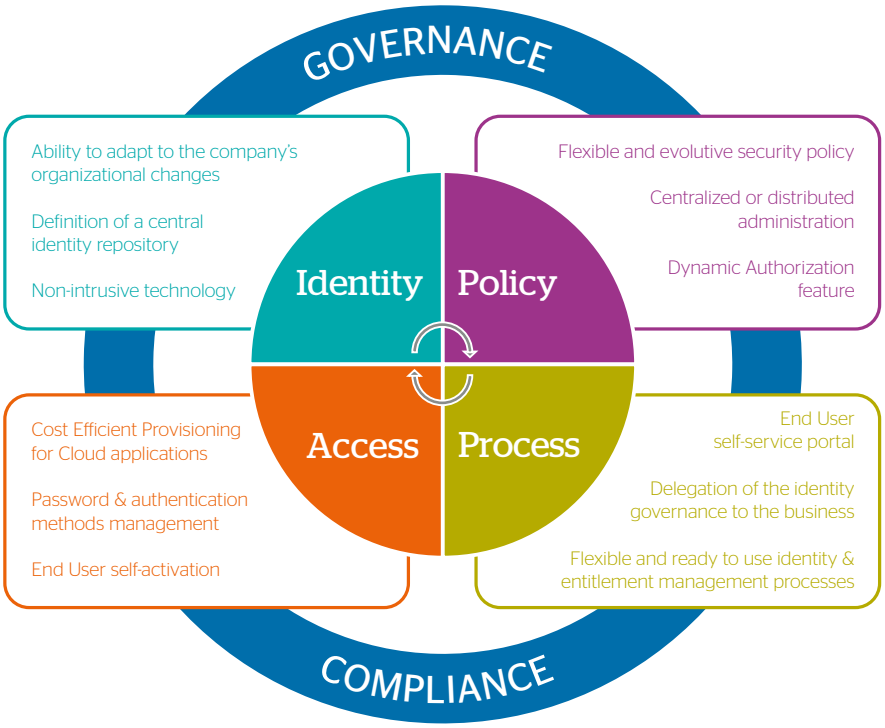
Update flows are multi-directional; therefore, each repository may be both client and information-provider-based, depending on the authoritative character of the data it contains.

An integrated upstream provisioning solution

There are two ways of using an identity-repository-creation solution: either independently and desynchronized from other identity and access management functions, or by integrating it into a complete solution. This second option allows you to immediately take account of the general problems and manage the 4 pillars of identity governance and administration:

- **Identities: :** first of all, identify and manage the users who can access the information system; not only employees but also subcontractors, partners, maybe even customers.
- **Policy: :** then, define and update the identified identity rights, and manage their evolution.
- **Processes:** allow your users to take part in the rights lifecycle management through the implementation of business-focused workflow processes.
- **Access: :** finally, apply and verify the security policy defined for identities and validated via processes on target systems. Allow each user to manage his accounts.

Evidian offers its **Evidian Identity Governance and Administration** solution to meet the requirements of the four pillars, as shown in the figure.



The ID Synchronization component facilitates the creation of the identity repository. This identity repository enables you to have an overview of all the identities that can access the information system, whether they represent employees, subcontractors, partners, or even customers if appropriate. These identities can come from different sources (Excel, CSV, SQL...), and can be created manually, or by the end-user himself with the self-registration feature. The communication between identity sources and the identity repository can be bidirectional: the simulation and the threshold management enable you to prevent the identity repository corruption. The identity uniqueness is guaranteed by the function that detects duplicates, even for identities during the retention period. It is also possible to declare users in advance by using a temporary identifier.

About Atos

Atos is a global leader in digital transformation with approximately 100,000 employees in 72 countries and annual revenue of around € 12 billion. The European number one in Big Data, Cybersecurity, High Performance Computing and Digital Workplace, The Group provides Cloud services, Infrastructure & Data Management, Business & Platform solutions, as well as transactional services through Worldline, the European leader in the payment industry. With its cutting-edge technologies, digital expertise and industry knowledge, Atos supports the digital transformation of its clients across various business sectors: Defense, Financial Services, Health, Manufacturing, Media, Energy & Utilities, Public sector, Retail, Telecommunications and Transportation. The Group is the Worldwide Information Technology Partner for the Olympic & Paralympic Games and operates under the brands Atos, Atos Consulting, Atos Worldgrid, Bull, Canopy, Unify and Worldline. Atos SE (Societas Europaea) is listed on the CAC40 Paris stock index.

Bull is the Atos brand for its technology products and software, which are today distributed in over 50 countries worldwide

Find out more about us
atos.net
ascent.atos.net

Let's start a discussion together

All trademarks are the property of their respective owners. Atos, the Atos logo, Atos Codex, Atos Consulting, Atos Worldgrid, Bull, Canopy, equensWorldline, Unify, Worldline and Zero Email are registered trademarks of the Atos group. Atos reserves the right to modify this document at any time without notice. Some offerings or parts of offerings described in this document may not be available locally. Pleasecontact your local Atos office for information regarding the offerings available in your country. This document does not represent a contractual commitment.
August 2017 © 2017 Atos

CT_170809_JG_Evidian-ID-Synchronization-WP-EN_01