

Safe Long-Term Archiving

Safe LTA



TINEXTA GROUP

TABLE OF CONTENTS

1	INFOCERT	4
1.1	CERTIFICATIONS	5
2	DESCRIPTION OF THE SOLUTION	6
2.1	ARCHITECTURAL SCENARIO	6
2.1.1	AMAZON OPENSEARCH	7
2.1.2	AMAZON S3	7
2.1.3	AMAZON KMS & SECRET MANAGER	7
2.1.4	SAFE LTA AS A MICROSERVICE	8
2.1.5	AUTHENTICATION WITH KONG AND KEYCLOAK	8
3	FUNCTIONALITIES	9
3.1	API FEATURES	9
3.2	USER INTERFACE	10
3.2.1	SEARCH AND RETRIEVE THE INGESTED AIPS	10
3.2.2	INSPECT A SINGLE AIP'S DATA AND METADATA	10
3.2.3	IDENTIFY AIPS WITH DIFFERENT LEVELS OF CONFIDENTIALITY	11
3.2.4	SIP INGESTION	11
3.3	MONITORING	12
4	MAINTENANCE AND ASSISTANCE SERVICE	13
4.1	CORRECTIVE MAINTENANCE	13
4.2	ADJUSTING MAINTENANCE	13
4.3	EVOLUTIONARY MAINTENANCE	13
4.4	HELP DESK OF SECOND LEVEL	14
4.5	SERVICE AVAILABILITY	14

INDEX OF FIGURES

FIGURE 1 SAFE LTA ARCHITECTURE	6
FIGURE 2 ADVANCED SEARCH.	10
FIGURE 3 AIP METADATA INSPECTION AND PACKAGE DOWNLOAD.	10
FIGURE 4 PACKAGE INGESTION.	11

PRIVACY

This document and the information it contains are confidential and proprietary of InfoCert, with the exception of information about the Customer.

The information contained herein is provided for the sole purpose of defining details relating to the relationship between the Customer and InfoCert and therefore may not be disclosed to third parties without the consent of InfoCert and the Customer.

1 INFOCERT

InfoCert S.p.A. is the largest Certification Authority in Europe and one of the main Qualified Trust Service Providers (QTSP) and is presents on the market as a highly specialized partner in digital Certification and electronic document management services. InfoCert can guarantee its customers the most innovative processes for managing their document and information assets.

With a share capital of 17,704,890 euros and a revenue of 72,9 million euros in 2019, InfoCert is a European leader in the processes for long time archiving of electronic documents and for certified e-mail services and is one of the largest Certification Authority in Italy for digital signature solutions.

InfoCert designs and develops paperless solutions with high technological value for document processes, through components of document management, long term archiving service, digital signature and certified e-mail solutions. Customers are followed in the choice of services and solutions that fully meet organizational and business needs, in compliance with general and sector-specific regulatory constraints.

InfoCert, with offices in Rome, Milan and Padua, is a qualified partner for companies operating in the banking, insurance, pharmaceutical, manufacturing, energy, utilities, commercial distribution, environment, quality, safety, health, public administration, trade associations and professional orders. InfoCert is also an 80% shareholder in Sixtema, the technology partner for the craft sector and SMEs in the CNA world, and 51% in AC Camerfirma SA, one of Spain's leading certification authorities.

Skilled professionals, with experience in the most modern technologies, and experts in Project Management, specialized in the customization and implementation of digital document management processes, give InfoCert a competitive advantage in the implementation of complex projects and solutions in the field of dematerialization. InfoCert's commitment to the development of cutting-edge solutions is underlined by the company's collaboration with some of the major universities and research centres in our country (Politecnico di Milano, EXO Organismo di ricerca, Università di Tor Vergata, Università di Salerno, SDA Bocconi). The ISO 9001, ISO 27001 and ISO 20000 quality certifications testify to InfoCert's desire to offer its customers the highest levels of service, also in terms of safety.

The models adopted by InfoCert S.p.A. give the customer a primary role: the full understanding of the needs and the design of customized solutions guarantee the achievement of excellence objectives.

InfoCert S.p.A. offers the market three components:

- **Consulting:** experience, dynamism and flexibility characterize the design and implementation of the most suitable solution for a completely digital management of documentation. InfoCert S.p.A. guides the customer in the transition from paper to digital management, optimizing document workflows and promoting the adoption of tools to support dematerialization, even in SaaS mode.
- **Technology:** modular, reliable and safe solutions with high quality technology and with full customer satisfaction. Continuous updates and important technological partnerships

guarantee InfoCert S.p.A. exclusive skills in design and development.

- **ASP services:** The InfoCert offer is divided into families of services, based on the latest generation of technology. InfoCert suites include services and solutions for the management of Certified Email (Legalmail), digital identity (InfoCert ID), digital certification and security (LegalCert), digital preservation of documents (LegalDoc, Safe LTA), integration of services in accordance with customer applications with cloud infrastructure (LegalCloud), electronic invoicing service (Legalinvoice), secure cloud storage solution (SecureDrive).

1.1 CERTIFICATIONS

InfoCert has the following certifications:

- **ISO 9001:2015**, is the Quality Management System aimed at achieving the company's objectives to ensure continuous improvement in meeting customer needs, optimize the organization of resources and interactions between business processes, reduce as much as possible the occurrence of situations and conditions of non-compliance of products and / or services. InfoCert's quality management system confirms the company's reliable structure, which guarantees the reproducibility of its performance, the maintenance and improvement of the quality standard of its services/products and guarantees the reliability of its production processes for customers, suppliers, employees and collaborators.
- **ISO 27001:2013**, is the Information Security Management System, certified for activities EA:33-35.
- **ISO 20000:2011**, is the Service Management System compliant with the international standard for IT Service Management, with the aim of maintaining and improving the alignment and quality of business services provided in relation to customer requirements, through a constant cycle of monitoring, reporting and review of agreed SLAs. The InfoCert Service Management System [SMS] model allows us to map and integrate the Service Levels (SLAs) guaranteed to customers in relation to the entire value chain of services [OLA and UC], facilitate the alignment between customer requirements and the InfoCert offer by setting up/defining formalized and measurable service agreements (SLAs) and guaranteed, ensure a control of the suppliers that contribute to the delivery of our services.
- **ISO 14001:2015**, is the Environmental Management System and responds to the company's strategy to implement a control of compliance with environmental regulations, an improvement of efficiency in processes, a careful response to customer and community requests with the aim of responding to responsible behaviour of the company.
- **ETSI EN 319 401**, Provision of Trust Services in accordance with Regulation (EU) 910/2014 eIDAS. With the entry into force of the eIDAS Regulation (01-07-2016 EU Regulation 910/2014) InfoCert has become Europe's leading Digital Trust Solutions Provider. For this reason, the company has been certified as a Qualified Trust Service Provider for trust services provided in accordance with Regulation (EU) 910/2014 eIDAS.

2 DESCRIPTION OF THE SOLUTION

Safe LTA is a cloud-based solution for the long-term archiving of digital documents. It is structured starting from the open RODA 3 platform (Repository of Authentic Digital Records), a long-term digital repository solution that delivers functionalities for all the main units of the OAIS reference model. Consequently, Safe LTA inherits RODA's compliance to standards such as the Open Archival Information System (OAIS), Metadata Encoding and Transmission Standard (METS), E-ARK Information Package specifications and PREMIS (Preservation Metadata).

Safe LTA is designed to be cloud-native on the AWS (Amazon Web Services) environment, adopting Amazon S3 (Simple Storage Service) for storing the ingested documents, and Amazon OpenSearch Service for indexing the related metadata.

Safe LTA is fully usable through its RESTful API. It also adds a level of security thanks to the Oauth2 standard to its API

Safe LTA can ingest E-ARK SIPs and transform them into AIPs according to the E-ARK specification by employing the *commons-IP* v2 Java package¹.

2.1 ARCHITECTURAL SCENARIO

Safe LTA maintains full compliance to the OAIS reference model (ISO 14721:2012). The information package (IP) modelling employed by Safe LTA is fully compliant with the E-ARK specifications.

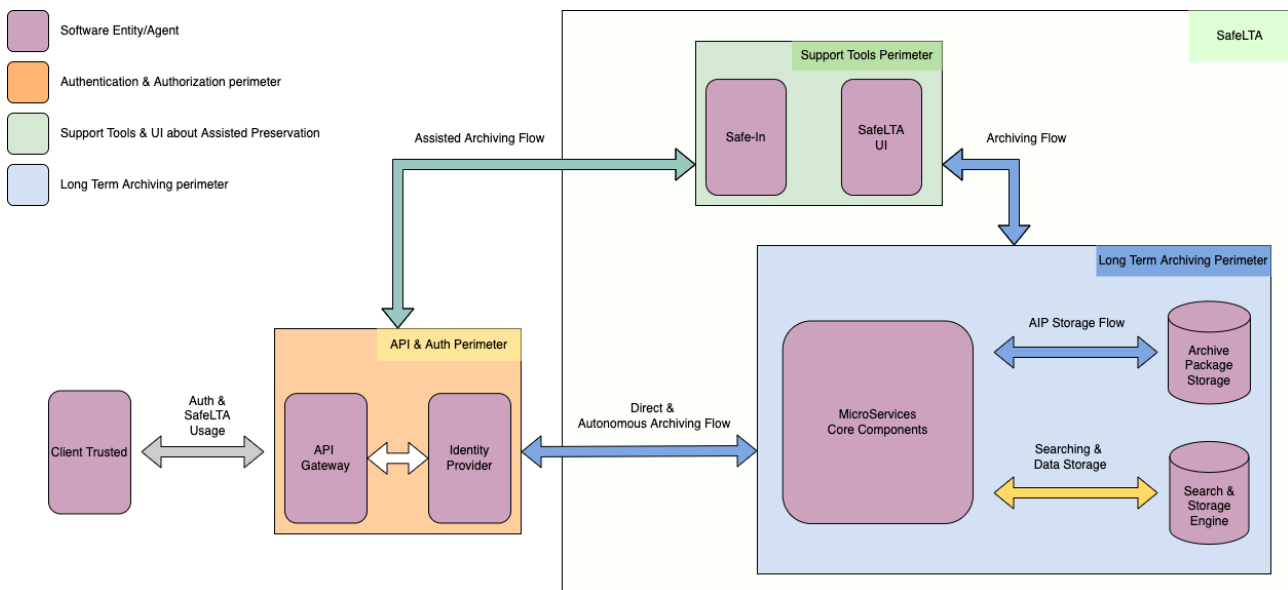


FIGURE 1 | SAFE LTA ARCHITECTURE.

¹ <https://github.com/keeps/commons-ip>

2.1.1 AMAZON OPENSEARCH

Safe LTA implements an abstraction layer on OpenSearch for indexing the ingested documents. OpenSearch is a fully managed service on AWS (namely Amazon OpenSearch Service), which greatly simplifies time-consuming cluster management tasks such as hardware provisioning, software patching, failure recovery, backups, and monitoring. Amazon OpenSearch Service allows users to store up to 3 PB (Petabytes) of data in a single cluster and can be easily scaled up or down via API calls or the AWS console. This service is also designed to be highly available using multi-Availability Zone deployments, allowing users to replicate data between three availability zones in the same AWS region.

The Amazon OpenSearch Service is also highly secure. Network isolation can be achieved with Amazon Virtual Private Clouds² (VPCs), and data encryption is carried out with AWS Key Management Service³ (KMS).

In case of a shard or node failure, OpenSearch can carry out cluster rebalancing automatically and rarely requires a manual intervention.

2.1.2 AMAZON S3

Safe LTA implements a new storage layer that provides native support for Amazon S3 (Simple Storage Service).

Amazon S3 is a storage service built to store and retrieve any amounts of data from anywhere on the Internet. It is a simple storage service that offers very high durability, availability, performance, security, and virtually unlimited scalability. The total volume of the data and the number of files that can be stored in Amazon S3 are unlimited, with only a size limit of 5 TB (Terabytes) per file. Files are temporarily cached locally to minimize the number of accesses to S3.

Access to Amazon S3 can be secured by employing adequate AWS IAM (Identity and Access Management) roles and by encrypting the content with the AWS Key Management Service (KMS).

2.1.3 AMAZON KMS & SECRET MANAGER

We provided integration with AWS KMS service and Secrets Manager to manage secrets objects used by Safe LTA, for example information related to eSeal Certificate used to sign some information about preservation (Mets files are signed and timestamped).

² <https://aws.amazon.com/vpc/>

³ <https://aws.amazon.com/kms/>

2.1.4 SAFE LTA AS A MICROSERVICE

With Safe LTA we also provide a first Microservices approach and in this way, we produced another component that simplify the current API-set and create for you a business API that introduce the concept of Document Class and SIP package generation for you. In this way you can define, in according to your specific legislation, the mandatory metadata associated with a document that needs preservation and the system will generate for you the SIP package and start the ingestion flow provided by Safe LTA.

2.1.5 AUTHENTICATION WITH KONG AND KEYCLOAK

Safe LTA exposes all its functionalities via its RESTful API constituted by a wide range of endpoints. Safe LTA exposes its RESTful API through Kong API Gateway⁴ and secures it with an access token authentication mechanism according to the OIDC standard.

We employ Keycloak⁵ as Identity and Access Management component used also for Identity provider federation where the Identity Provider are OIDC or SAML compliance. When a new user is added to Safe LTA, his/her information will be immediately available to Keycloak. The current *username* value is included in the access JWT (JSON Web Token) returned by Keycloak upon successful login. Also, for client integration about backend-to-backend interaction, we use Keycloak to manage it.

After validating the token, the API Gateway extracts the *username* from the token and adds it to the request forwarded to Safe LTA. The authentication is completely delegated to Keycloak. Authorization, on the other hand, is completely handled by Safe LTA since each call to its API includes the corresponding *username* in a dedicated HTTP request header.

⁴ <https://konghq.com/kong/>

⁵ <https://www.keycloak.org/>

3 FUNCTIONALITIES

3.1 API FEATURES

Safe LTA is a cloud-native application that can be effectively and efficiently scaled up or down automatically by Kubernetes depending on the current workload of the system. This architecture, backed up by Amazon's Elastic Compute Cloud (EC2), ensures the availability of the system's functionalities under any circumstances and transparently to the end user, without human intervention.

Safe LTA can be easily integrated with other systems through its RESTful API. They can be used to perform a wide range of functionalities, including:

- Provisioning
- Users, groups and permissions management
- SIP ingestion and transformation into E-ARK AIP
- Advanced search operations
- Document and metadata retrieval
- Download of the AIP as ZIP archive

Safe LTA not only validates standardised SIPs (submission information packages), but also handles file format identification, extracts technical metadata, and migrates file formats to more "durable" alternatives (PDF only). The authenticity of the ingested data is ensured by recording PREMIS metadata every time an action is performed on a digital object.

All interactions between users and the repository (human and software) are logged for security and accountability reasons.

Each endpoint is secured according to previous paragraph "*Authentication with Kong and Keycloak*".

Documentation is available on <https://developers.infocert.digital/>.

3.2 USER INTERFACE

3.2.1 SEARCH AND RETRIEVE THE INGESTED AIPS

Logged users will be able to search and retrieve the ingested AIPs based on a given set of metadata. The users will be able to perform full-text searches over one or more indexed terms.

Users will be able to perform date-range queries on date fields, exact term match and full-text match queries on text fields. Advanced search operations where two or more search clauses are combined will also be possible (see FIGURE 2).

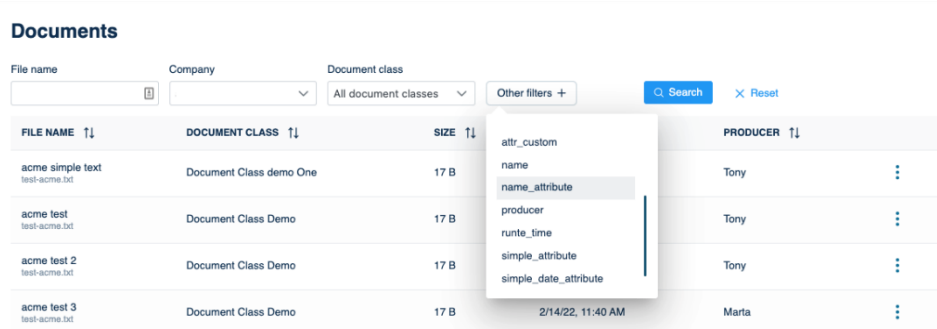


FIGURE 2 | ADVANCED SEARCH.

3.2.2 INSPECT A SINGLE AIP'S DATA AND METADATA

Users will be able to inspect each AIP, view its representations (data files, metadata see FIGURE 3). Moreover, users will also be able to retrieve the current AIP as a ZIP archive file.

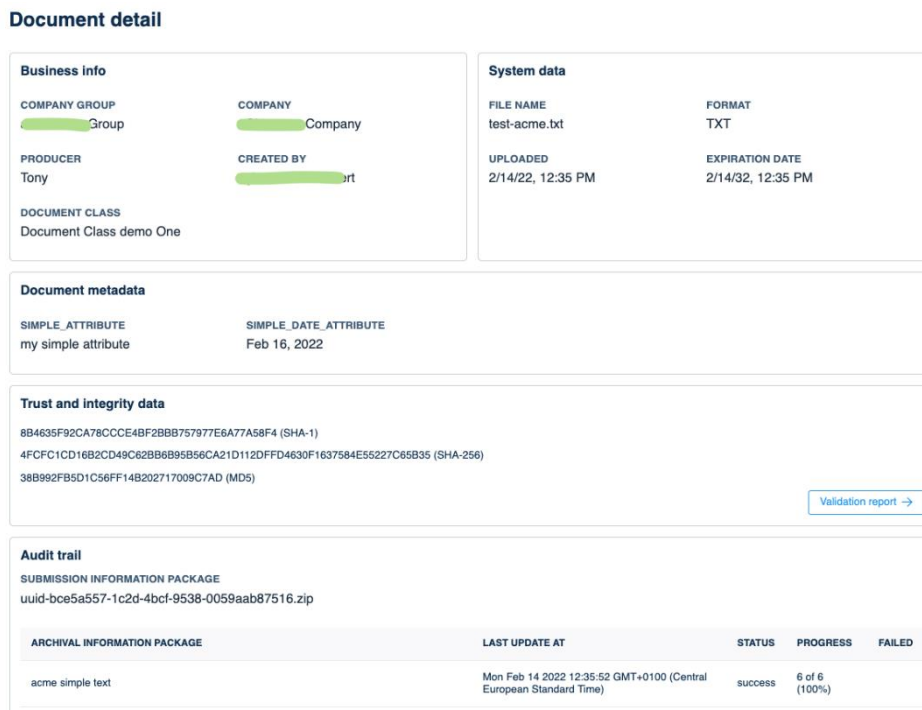


FIGURE 3 | AIP METADATA INSPECTION AND PACKAGE DOWNLOAD.

3.2.3 IDENTIFY AIPS WITH DIFFERENT LEVELS OF CONFIDENTIALITY

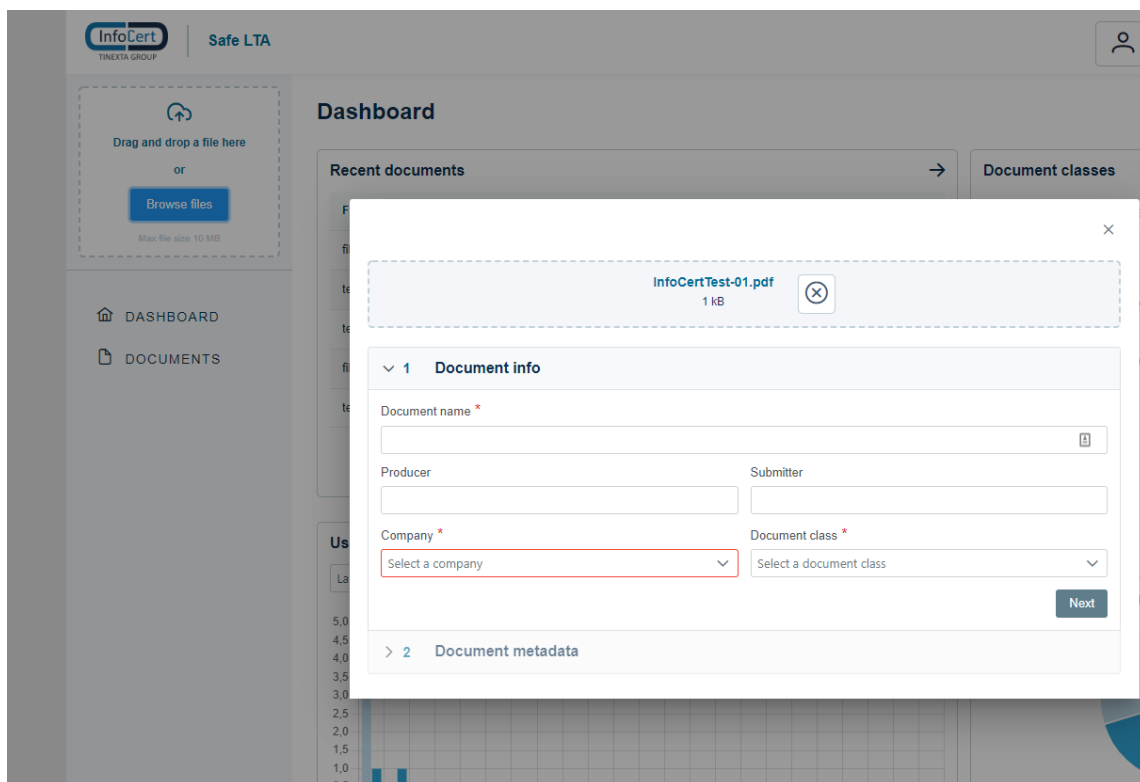
The access level to the AIP packages is in accordance with the accessibility levels of the roles provided by Safe.

A user with adequate permissions and group memberships could easily identify AIPs with a given level of confidentiality by searching for the AIPs belonging to the corresponding user-group. The AIP permissions mechanism grants the correct segregation of documents based on group-wise permission settings.

3.2.4 SIP INGESTION

Logged users will be able to ingest SIP package.

To perform file preservation all mandatory data must be provided: file name, producer, mime type and search indexes.



The screenshot displays the InfoCert dashboard interface. On the left, there is a sidebar with navigation options: 'DASHBOARD' and 'DOCUMENTS'. The main area shows a 'Dashboard' header and a 'Recent documents' section. A modal window is open, titled 'InfoCertTest-01.pdf' (1 kB). The modal contains a form for document ingestion with the following fields:

- Document name * (text input)
- Producer (text input)
- Submitter (text input)
- Company * (dropdown menu with 'Select a company' selected)
- Document class * (dropdown menu with 'Select a document class' selected)

A 'Next' button is located at the bottom right of the form. Below the form, there is a section for 'Document metadata' which is currently collapsed.

FIGURE 4 | PACKAGE INGESTION

3.3 MONITORING

Safe LTA is under monitoring by Infocert for two main aspects:

1. Safe LTA application data log about ingestion and preservation flow collected on file system and ingested by NewRelic
2. Infrastructure with data collected by AWS CloudWatch and NewRelic

4 MAINTENANCE AND ASSISTANCE SERVICE

As part of the Maintenance and Support service, InfoCert undertakes to maintain or restore the components of the services provided to good working order.

The interventions for the entire duration of the contract, are related to three types:

- Corrective maintenance
- Adjusting maintenance
- Evolutionary maintenance

All reports must be received through the 2nd level Help Desk service (described below) which will be provided through the web ticketing system provided by InfoCert.

All interventions will normally be carried out remotely.

4.1 CORRECTIVE MAINTENANCE

The service is designed to eliminate anomalies detected during operation and which are present in the components of the services provided.

Corrective maintenance operations do not modify or extend existing functions or the original application architecture but leave the size of the baseline unchanged.

The process of managing correction requests is triggered by the second level support provided by the Help Desk, described below.

The primary objective is to ensure that any application malfunction is resolved effectively and promptly, so as not to cause damage to the impacted business functions.

The process starts with the notification of a user, proceeds with the opening of a ticket and ends with the resolution of the anomalies and any updating of the supporting documentation.

We will always take care to verify possible opportunities for preventive maintenance that permanently eliminate the causes of the malfunction. To release the correction, the necessary checks and regressive tests will be carried out.

4.2 ADJUSTING MAINTENANCE

The purpose of this service is to update the software components supplied with the services provided to ensure that they are adapted to any changes in current national/European legislation.

The time required for an adjustment will be agreed between InfoCert and the customer.

4.3 EVOLUTIONARY MAINTENANCE

On the unilateral initiative of InfoCert, new functionalities or modifications to existing functionalities may also be implemented and made available, in order to maintain the quality level of the product and its adherence to the technological and regulatory standards of the market.

The service does not include evolutionary, technological and functional interventions, which are explicitly requested by the Customer. In this case, InfoCert is available to carry out the technical and

economic evaluation of the requested interventions and to discuss the relative supply as an extension of the service.

4.4 HELP DESK OF SECOND LEVEL

InfoCert offers the contact service to the "Key Users", professional figures of the Customer who have the task of collecting and interpreting the needs of the various internal users and who represent the privileged interlocutors in the relationship between the Customer and InfoCert.

The single claim is managed by the staff of the Help Desk performing the activities of:

- Acceptance and recording of Users' calls
- Classification of the event, with resolution of false positives or, if not, opening of a ticket with the description of the claim
- If possible, provision of a solution in the case of "known issues" or provision of any "workaround" waiting for a final solution and activation of the escalation procedure
- Control of the activated escalation procedures, checking the results and closing them
- Documentation of the entire process of the assistance service

Analysis of statistics on interventions, to identify needs and define problem prevention actions.

4.5 SERVICE AVAILABILITY

The assistance and maintenance services are guaranteed with the following working hours:

- from Monday to Friday (excluding national and midweek holidays).
- from 09:00 AM to 6:00 PM.

For the Safe LTA service, InfoCert guarantees a 99% service availability on a monthly basis.