



eDAMIS – the application which enables the transmission of statistics from Member States to Eurostat

By John Allen and Vincent Tronet (ESTAT B3)

eDAMIS (electronic Data files Administration and Management Information System) is a remarkable example of a successful working application which implements the eGovernment vision to support the interaction of a Commission DG – Eurostat – with its counterparts in the national administrations of Member States.

eDAMIS is The Single Entry Point for statistics at Eurostat.

Eurostat is the Commission DG responsible for implementing EU policy on statistics. EU Member States regularly transmit statistics according to more than 150 legal acts (as well as informal agreements and ad hoc projects) - an estimated total of around 36000 “dataset occurrences¹” per year. Around half of the data are sent by National Statistical Institutes, the rest come from nearly other 1000 national authorities which are responsible for statistics in specific domains.

The Single Entry Point (SEP) concept is that data for all statistical domains should arrive at a common reception area in Eurostat, so they can be automatically monitored, checked and delivered into the target production environment, with a set of common informatics tools. The SEP implies that incoming data files are identified as being instances of a dataset which is included in the inventory of datasets to be transmitted by Member States to Eurostat. The aim of this activity is to ensure efficient, secure and monitored transmission of statistical data from Member States to Eurostat, as the first step in an efficient in-house data life cycle (CVD²). Eurostat’s ongoing CVD project

aims at rationalisation of the whole data life cycle, by increasing automation and eliminating parallel IT systems at each stage from data reception to publication on the Eurostat [web site](#) .

The SEP is an organisational concept which is implemented through eDAMIS, an integrated family of IT applications, which assure secure transmission of data files, as well as monitoring and delivery to production units.

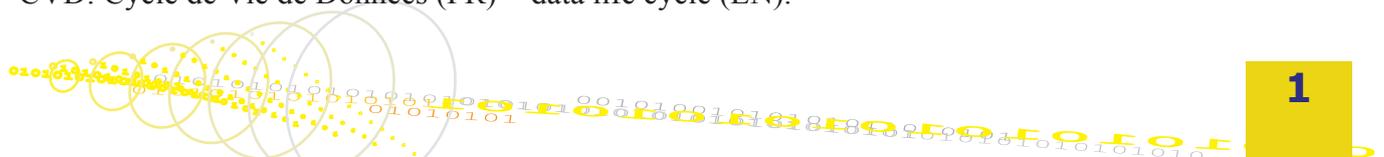
History

In the dark ages - indeed as recently as 1995 - many data arrived in Eurostat on paper, tapes, cartridges and diskettes (Figure 1).



¹“dataset occurrence”: an occurrence of a dataset for 1 country for 1 period (or time series or sequence), excluding replaced and appended files. Since for various reasons it is often necessary to send revised data, the number of data files transmitted is around twice the number of dataset occurrences.

²CVD: Cycle de Vie de Données (FR) = data life cycle (EN).



Then email came and took over as the preferred means of transmission. However, data transmission by email carries a number of risks for Eurostat and is very dependent on human beings. If personal mailboxes are used, data may not arrive to the correct person, or not at all, if he or she is on holiday or has left the job.

In view of the fact that 80% of incoming emails may be spam or contain viruses, there is a significant chance email messages carrying statistical data will be blocked in the Commission's email filtering system. There is also a risk of disclosing or misdirecting confidential or sensitive data if transmission is done by email.

Reception of data by email also creates a heavy workload for Eurostat's production units. In a survey of data treatment practices, it was found that the equivalent of 72 full-time staff are engaged in receiving data, preparing them and loading them into production databases. This is about one-third of the statistical staff in the production units.

With the SEP, the risks linked to email transmission and the workload for units can be significantly decreased. Data are automatically routed to the correct production environment. Receipt notifications are sent out automatically and all partners involved in data transmission can also consult monitoring reports which show which countries have delivered what and when. Apart from monitoring, the receipt notifications and the monitoring reports are valuable from a legal perspective. Countries can see if they have followed their legal obligations and delivered data in time.

eDAMIS components

eDAMIS incorporated several pre-existing tools (Stadium, Statel, Ediflow, Genedi) while adding new functionality, notably the possibility to upload data files to Eurostat via the Internet.

The present-day architecture of eDAMIS is shown in Figure 2. at the end of this article.

From the perspective of Member States, there are two visible components of eDAMIS:

eDAMIS Web Application (eWA): NSIs normally send data through eWA. This application is installed on one server at the NSI, and can be accessed by authorised NSI staff members on their intranet, through a web browser interface. eWA can also be used for fully automatic transmission of data files to Eurostat.

eDAMIS Web Portal (eWP): Eurostat receives data from around 1000 different national authorities. Unlike NSIs, many of these send small numbers of data files to Eurostat, so that they can not be expected to use applications which require a local installation. eWP is a [web portal](#) through which data files can be uploaded to Eurostat (see the screenshot in Figure 3 at the end of this article).

This is a simple solution for national authorities other than NSIs that currently send data by email. eWP requires no local installation as it is used via a normal Internet connection and a web browser.

In the latest release (eDAMIS 2.5), eDAMIS also provides a "web form" interface, which allows users in Member States to prepare and check a dataset in a screen form before uploading it to Eurostat.

Operation of the eDAMIS service inside each NSI is supported by a so-called Local Coordinator for data transmission. Eurostat provides continuous technical support for the Local Coordinators. It should be noted that for Local Coordinators at NSIs, the portal (eWP) provides access to other functions of the eDAMIS environment, such as the inventory of datasets to be sent to Eurostat and the management information system, as well as user management and control of user access rights.

The Management Information System (MIS), accessed through the eDAMIS portal, gives access to online data traffic reports, which are updated in real time and show the actual reception date as well as the indicative deadline for each dataset occurrence.

There is an extensive online help system, the [eDAMIS Help Centre](#) , built in CIRCA.

All normal user functions of eDAMIS are accessible





through a browser. Invisible to normal users, but a vital part of eDAMIS, are two more components:

The **eDAMIS/Stadium server** and the **eDAMIS Kernel** (which is part of eWP) constitute the central hub of the eDAMIS system. The original Stadium server has been in service for many years; the eDAMIS Kernel was developed since 2003. Together, these provide many back-office services for eDAMIS, including: the inventory of datasets transmitted to Eurostat by Member States, the reception and delivery of data files; the generation of notifications (to users receiving data), acknowledgements (to users who have sent data) and reminders (to users who are expected to send data); and the storage and query of information for the MIS.

STATEL is the communications “middleware” for eDAMIS, which has been in use since 1991. It creates a mirrored “Virtual File System” (VFS) linking components of eDAMIS (currently eWA) installed in Member States with the eDAMIS/Stadium server. It transfers data files in segmented packets, with encryption, and ensures that all segments of a data file are transmitted successfully. It was designed to be independent of the data transmission (communications) protocol and the host operating system. Depending on

the Member State involved, transmission via STATEL can use the FTP or HTTP protocol (SMTP is also possible, but is no longer used), and there are Unix/Linux and Windows versions.

In fact, Statel can be used for transmission of data to the eDAMIS/Stadium server without the use of the eDAMIS application (eWA).

DG Agriculture also uses Statel, together with its own implementation of eWA and the Stadium server.

Languages and platforms

The components of eDAMIS were developed with different languages and tools, for use on multiple platforms, as summarised in Table 1.

Security

As many statistical data transmitted to Eurostat are legally confidential, and because it is generally important to avoid accidental disclosure of any statistics which have not been cleared for publication, eDAMIS implements security measures of several kinds.

1. STATEL identifies sending and receiving entities by so-called Statel Nick Names (SNNs) using a specific mechanism of passwords. These STATEL names are not system accounts. Even if their passwords were broken, this would not give any rights on the operating system where STATEL is installed.

Table 1: Languages, application environments and operating systems

Component	Langages/application environment	Operating systems
eDamis kernel	ColdFusion, Java, Oracle	Unix
eDamis Web Portal	ColdFusion, Java, Oracle	Unix
eDamis Web Application	Java	Unix/linux, Windows
Stadium server	ColdFusion, Oracle, Perl, C++	Unix
Statel	C	Unix/Linux

2. All data transmitted via STATEL are encrypted. This encryption is automatic and requires no special action by the sending user. Starting with STATEL 3.6, a better encryption algorithm was introduced (a symmetric Vernam cipher, using a seed based on characteristics of the file, such as the STATEL Nick Names, the file name and the local



sequence number. The key which is dynamically created is 256 bits long.). This means that obtaining and installing a copy of STATEL no longer suffices to decrypt intercepted STATEL data packets, known as PDUs (Protocol Data Units), but that the complete environment of the targeted SNN needs to be reconstructed. This encryption method, though safer than the original method, is not unbreakable. The current key is computed from static data, and hence it is possible to reconstruct the key when this static data is available (as it is in the PDU or the permanent STATEL data structures on disk.)

The latest versions 4.x of STATEL support full-strength encryption mechanism based on openssl, with Diffie-Hellman key agreement upon reset (length of keys: 512 bits), and 3DES encrypted PDU payload (length of key: 192 bits), ensuring excellent security in transit. The STATEL 4.x key exchange mechanism is based on asymmetric keys (public and private keys), but the encryption of the data is based on a symmetric key. This approach (combined usage of asymmetric and symmetric keys) is comparable to SSL or https. It is considered as highly secured, but also performant for data encryption/decryption (a lot more performant than the usage of asymmetric keys for encrypting/decrypting data).

The use of segmented data packets (PDUs) in combination with SNNs and encryption, means that data transmission using STATEL from eWA in the Member State to the Stadium server at Eurostat may be considered highly secure.

3. To provide an additional level of security for confidential datasets throughout the transmission chain from eWA up to delivery in the secure production environment, eDAMIS also offers the option of asymmetric encryption; this requires preparatory action by the domain managers concerned.
4. For uploading data files through eWP, an ht-

tps connection (using SSL) is always used, and asymmetric encryption is an option which requires preparatory action by the domain managers concerned.

5. In relation to (3) and (4), there is a formal recommendation to units and domain managers responsible for confidential datasets that they should carry out the necessary preparatory actions to set up asymmetric encryption. In practice, this recommendation is not always followed.
6. User authentication and access rights for eWA and eWP use a common system based on the use of the CIRCA LDAP.

Throughout the implementation of eDAMIS, Eurostat has followed the principle that security measures should as far as possible be built in to the system and require as little action as possible by users. Even without the use of the asymmetric encryption option – which is recommended for confidential datasets, but not always used – there is very little risk of unauthorised disclosure of statistical data through the use of eDAMIS, and it is certainly more secure than any alternative mode of transmission (email, physical delivery...). Up to now, there is no known case of unauthorised disclosure of statistical data during transmission via eDAMIS.

HOW DO WE MEASURE THE SUCCESS OF eDAMIS?

We use one main indicator to measure the success of eDAMIS, which is the number of first transmissions of dataset occurrences as a percentage of the total number of expected dataset occurrences, over a period of a year (or in certain cases a 3-month period). This “Single Entry Point coverage indicator” has grown from 18% for 2003 to 33% for 2006.

In the first quarter of 2007, Eurostat received an average of about 50 new dataset occurrences each working day. At the end of a quarter, when many statistics





are delivered, the traffic level can be much higher: for example, 576 data files were received in one single day, on Friday 30/6/2006 (last day of the week, last day of the quarter and just before holidays).

We have operational targets for the SEP coverage indicator in the current and future years. Currently these targets are: 2007: 60%; 2008: 70%; 2009: 90%. We recognise that these targets are ambitious, but they help us to translate the decisions of top management at Eurostat and in NSIs into reality. In fact, several NSIs already send a very high proportion (possibly >90%) of their data through eDAMIS; it is much more difficult to achieve high usage of eDAMIS by the numerous national authorities other than NSIs, which individually send only small numbers of data files. The introduction of eWP in 2006, with the possibility of uploading files through the portal with no local installation, will encourage non-NSI national authorities to use eDAMIS.

Another indicator of practical significance is the number of NSIs which have installed eWA. The first beta version of eWA 1.0 was installed in 2 NSIs at the end of 2004. As of April 2007, the NSIs of all EU-27 NSIs had working eWA installations. eWA is also running in the NSIs of the EEA/EFTA countries (except Liechtenstein), the three candidate countries and Serbia.

WHAT FACTORS HAVE CONTRIBUTED TO THE SUCCESS OF eDAMIS?

Identifying some factors which have contributed to the success of eDAMIS may help other DGs which are building IT applications to support their regular business with national administrations. We have identified the following factors of success:

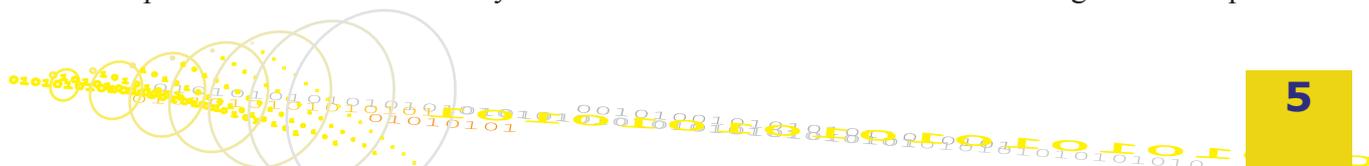
Simplicity: from the perspective of the ordinary user – whether in a national authority or at Eurostat – the everyday use of eDAMIS demands no more effort than the use of email, which is in reality the main alternative to eDAMIS. Through successive versions, the interface and the sequence of steps needed for the basic operations have been steadily refined to the

point that users need almost no training on the application itself, once they understand what it does for them. For example, users have told us that they like to send and receive data by email because it is easy to add some explanations in the email or in an attached document. Therefore eDAMIS allows both for short text annotations and for attached documents with the sending of any data file.

Time: although there has been a phase of intensive development work on eDAMIS since 2003, it builds on applications which were conceived by Eurostat as long ago as the 1980s. The present-day eDAMIS is not a single application, but an “integrated environment” of applications and modules some of which have been in use for at least 10 years. This length of time has contributed not only to the maturity of eDAMIS from an IT point of view, but also to its organisational acceptance within Eurostat and within national authorities; it is seen as something stable, which will still be around five or ten years from now. Most remarkably, it has survived several major reorganisations at Eurostat and an almost complete change in the team of people responsible – although one colleague has worked on eDAMIS and its predecessors since the 1980s.

Modularity: this means that development work and the addition of new functionality has been concentrated on some parts of the eDAMIS environment, which in turn can make use of stable, reliable code from the older elements of the system. When the present phase of intensive development of new functionality slows down in 2008, then more attention can be given to re-writing the older parts of eDAMIS in the interests of ensuring their future stability and maintainability.

Use of existing services: not only does eDAMIS use some components which are in effect separate applications – most notably STATEL – but it also uses services provided by other Commission applications. The most creative example of this is that registration and authentication of users in national authorities is done via CIRCA. This means that information on users in national authorities is maintained in the CIRCA LDAP. Eurostat was the original developer of CIRCA



and many staff working in NSIs and other national authorities already have CIRCA userids which they use to access documents for Eurostat committees and working group. These users do not need to re-register for eDAMIS. Information on Eurostat users is taken from COMREF and they are authenticated through the normal Commission LDAP.

Local Coordinators: use of eDAMIS in NSIs is supported through a network of Local Coordinators (LCO). The organisational role and hierarchical status of the LCOs varies widely from country to country, but in general they provide support for the local eDAMIS installation - eWA - as well as acting as a communication channel between Eurostat and the NSI users. They also intervene in the granting of specific user access rights to NSI users. LCOs meet twice a year in Luxembourg at the Local Coordinator Working Group (GLC). There is also a monthly newsletter for LCOs - the Data Transmission Courier - currently up to issue 59. LCOs are nominated by NSI management and the job has changed hands frequently in some Member States, but there is a stable core of very experienced LCOs who use the GLC to transmit their opinions about the operation of eDAMIS and the ways in which the applications could be improved.

Domain Managers (staff responsible for specific statistical domains at Eurostat) are the key in-house users of eDAMIS. They are in charge of updating the inventory of datasets linked to their domains and of granting rights to users in Member States (for some NSIs, in collaboration with the LCO). They also play a leading role in promoting the usage of eDAMIS for their data senders, in organisations other than NSIs.

User feedback: in addition to the GLC, LCOs as well as ordinary users are encouraged to submit bug reports and requests for new features through an online “change request” (RFC) [form](#) accessed through a menu in the eDAMIS Help Centre. RFCs are visible to all users, together with information on the follow-up by Eurostat. For the main eDAMIS components (eWP and eWA) the development road map foresees

major releases at approximately six-month intervals, so it is often possible to tell users that a change will be implemented in one of the next major releases.

Support team: a three-person helpdesk provides first-level support for eDAMIS not only on normal working days but also on days when the Commission is on holiday but work continues normally in national authorities. The helpdesk team also carry out system maintenance operations and contribute to the development of eDAMIS. In addition to the helpdesk, data transmission support team comprises four other Eurostat staff in the section “data transmission tools and services” plus other unit staff who regularly participate in team meetings or provide expertise on specific issues. There is also an external support team - with the equivalent of four more experts - run by a contractor, which provides backup to the in-house team as well as carrying out specific support tasks such as writing manuals and providing training courses.

Money: it goes without saying that eDAMIS has cost - and continues to cost - a large amount of money. Much of the early development was funded through the IDA-I and IDA-II programmes, up to 2004. Since 2005, development and support costs - running at around 1 mio euro/year - have been funded mainly through the Eurostat operational budget, with a small part of the support costs plus the infrastructure costs funded from the Commission IT budget. No IDA-BC funding has been sought for eDAMIS, as it was thought that this would go against the IDABC management committee’s intention that DGs should seek durable solutions for corporate applications rather than relying on programme financing.

FUTURE ISSUES

In order to support the requirements of users, the functionality of eDAMIS will have to increase over time; we have seen that as users gain experience with the application, they are highly creative in suggesting new things that eDAMIS ought to do. This is particularly true of the more experienced Local Coordinators in NSIs and domain managers at Eurostat. At the



same time, it is desirable to rewrite elements of the system to improve stability, performances and maintainability, and to consider whether new technologies can be exploited.

From an IT perspective, four possible areas of future development are of special interest. The first concerns the complete integration of the original Stadium server with the eDAMIS Kernel. This has been under consideration since the start of development of eWP and the Kernel, but due to the critical functions of Stadium within the system, and the reliability of Stadium as a mature component in production since 1993, it was decided to link these two elements rather than to merge them completely. However, the development road map includes the target of fully integrating Stadium within eDAMIS in eDAMIS 3.0, scheduled for December 2008.

The second concerns the use of the standards and architecture of the SDMX initiative³. The acceptance of the SDMX approach in the European Statistical System will lead to a much greater use of the SDMX XML standard (SDMX-ML) not only for data transmission but also for transferring data (and metadata) between Eurostat applications. For certain types of data, the existing “push” transmission will be replaced or supplemented by the “pull” approach, where NSIs make datasets available via web services and these data are retrieved automatically by a Eurostat application. This architecture has already been tested within the IDABC X-DIS⁴ project and is being integrated into the eDAMIS architecture during 2007-2008;

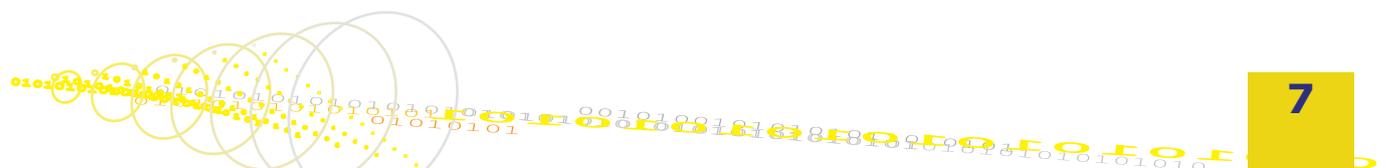
Third, with the replacement of CIRCA by CIRCABC during 2007-2009, it will be necessary to change the

user authentication system of eDAMIS. We assume that eDAMIS will move to the use of ECAS for Eurostat users. For users in national authorities, we will follow the approach used for CIRCABC. The scheduling of the changes in eDAMIS has not yet been determined, as we await developments in CIRCABC.

Fourth, it is possible that eDAMIS itself might offer web services interfaces for data reception and other functions. Web services are already used when linking eDAMIS to CIRCA (for user authentication) and by the SDMX applications. In 2006-2007, the eDAMIS team made a small feasibility study to explore the technology for web services in data transmission and develop a small demonstrator application. The eDAMIS team has also been involved in the IDABC Interoperability Expert Group which is looking at the definition of a web service profile for eGovernment services. In 2007-2008, it is intended to launch a small pilot project on possible web services in eDAMIS with one or two Member States. However, the rate of progress in this area will depend on decisions by DIGIT as to whether to permit and to support external web services

³ SDMX: Statistical Data and Metadata eXchange. The SDMX initiative is run by a group of seven sponsor organisations (Eurostat, the Bank for International Settlements (BIS), the European Central Bank (ECB), the International Monetary Fund (IMF), the Organisation for Economic Co-operation and Development (OECD) and the United Nations Statistical Division (UNSD) and the World Bank) which have joined together to de-

⁴ X-DIS: XML for Data Interoperability in Statistics



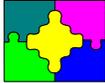


Figure 2: eDAMIS architecture in 2006

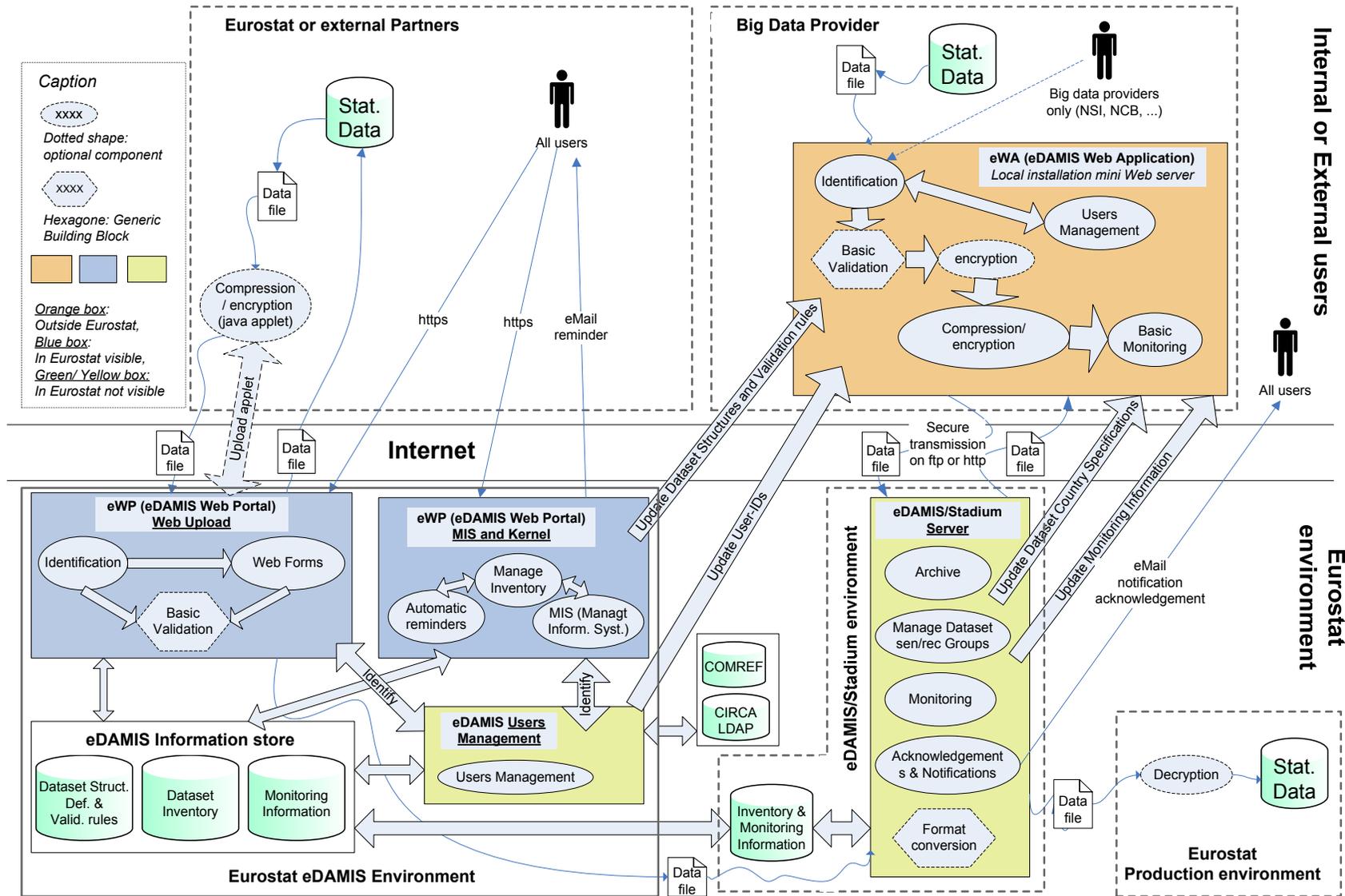


Fig 3: eDAMIS Web Portal Upload screen

