E-MONITORING AND E-REPORTING WORKSHOP

Pacific Islands Forum Fisheries Agency Headquarters, 31 March – 1 April 2014
Honiara, SOLOMON ISLANDS

ELECTRONIC REPORTING IN FISHERIES

WCPFC-EmandErW-2014-DP01
29 March 2014

A paper by the European Union

EUROPEAN COMMISSION
DIRECTORATE-GENERAL FOR MARITIME AFFAIRS AND FISHERIES

MEDITERRANEAN AND BLACK SEA
Integrated Fisheries Data Management

Brussels,
MAre D4/Fy

Subject: Electronic reporting in fisheries

1. INTRODUCTION

The intended readership of this document is anyone who needs a general overview of the on-going development of a data exchange network for fisheries. It reflects the state of play on 27/03/2014.

Some of the content is somewhat technical but an effort has been done to avoid deep technical explanations. A lot of documentation is available upon request.

A brief technical introduction to some of the technical terminology is annexed.

2. BACKGROUND

Data, as basis for analysis and decision making, is gaining importance in most sectors and fisheries is not escaping from this trend.

As a consequence, most fishing nations, and related organisations, are developing electronic reporting and data exchange systems to improve fisheries monitoring and control, to extend the data availability for scientific research, combat illegal fishing and ensure health and safety.

A good overview of initiatives and approaches is available in the "Potential for E-Reporting and E-Monitoring in the Western and Central Pacific Tuna Fisheries" document made for WCPFC.

Naturally, each coastal state, or RFMO, expects that its regulations and data requirements are met by all vessels (including foreign ones) that are fishing in their area.

The risk is that, without harmonisation of data exchange requirements, procedures and related IT systems of the fishing nations and RFMO, each country, and its fleet, will need a range of IT systems and data exchange methods depending on where they are fishing. The cost of this could strongly delay progress.

Commission européenne/Europese Commissie, 1049 Bruxelles/Brussel, BELGIQUE/BELGIË - Tel. +32 22991111
Office: J-79 - Tel. direct line +32 229-64013

francky.callewaert@ec.europa.eu

The European Union (EU) has mainly experience, and is even partly at the origin of:

- The exchange of vessel position data (VMS) based on the North Atlantic (NAF) format

- The exchange of catch on entry, catch on exit and catch messages within the context of the NEAFC and NAFO, and also using NAF

- The NOR-ERS electronic logbook (also called CREWS) for data exchanges between Norway and EU for fishing activities in each other's waters.

- The EU Internal EU-ERS (version 3.1) electronic logbook, mainly used for fisheries in EU waters.

Although the merits of all those systems must certainly be recognised, not in the least because some of these systems result from ground breaking work and are still very much in use today, it is also very clear that their shortcomings prohibit smoothly spreading the use of electronic reporting, while their "business content" can't keep pace with the fast growing need to exchange more and more different data.

Indeed, next to data immediately related to on-going fishing activities (VMS, catch data) for control purposes there is a growing need to e.g. exchange vessel and licence data, aggregated catch data, and scientific data.

Further, requirements for data exchange with other communities like port authorities, coast guard, food and drug administrations, auctions … are being formulated.

The effect is that fishing nations are confronted with a growing number of data management and data exchange requirements, each leading to significant IT development and maintenance costs.

Those costs, not to mention those for training of staff in using those systems, become a major barrier for expanding the use of electronic systems, hence also for improving fisheries management.

3. ALTERNATIVE APPROACH

To overcome the above described hurdles, the EU (DG Mare) has started the Integrated Fisheries Data Management (IFDM) programme. Within this context a general approach was formulated, common to all individual projects under the programme, which should lead to a high quality data exchange environment.

The criteria are the following:

Performance:

- Easily connecting all stakeholders adopting the system
- Efficient data storage and exchange respecting the role and requirements of stakeholders

Flexibility:

- Able to adapt to the different requirements of all the various stakeholders,
- Fast introduction of new needs
- Easy resolution of occurring shortcomings and errors

Cost:

- Affordable for all stakeholders, including smaller scale fisheries
- Able to connect to existing IT systems

The tools to achieve this are:

- Harmonisation of business practices,
- Standardisation of workflows and data requirements,
- The gradual construction of a modular "open source" IT platform available to all parties.

The above does NOT mean imposing identical systems everywhere. It is about creating a common framework to which each party is invited to contribute but also respecting individual circumstances.

For cooperating parties managing fisheries in a particular area, it means integrating their needs in the existing framework. For cooperating fishing nations it means expanding their already existing platform with existing or new modules depending on where they are fishing.

The construction of this has started with the FLUX Project. This project builds a business independent data transportation layer called FLUX Transportation Layer (can be understood as a secure email system between computer systems), and is standardising business content according to UN/CEFACT standardisation.

4. FLUX TRANSPORTATION LAYER (FLUX TL)

4.1. Introduction

One major criticism towards most existing data exchange systems is that business content and data exchange technology are usually closely coupled. Over time, and for the data exchange priority of the day, a different data exchange mechanism has been produced.

Today, the EU is confronted with data formats varying from NAF, to CSV and PDF Forms, up to XML and with data exchange mechanisms from floppy disks, over email to web services. The costs for maintaining that diverse technology park are enormous.

One of the first ambitions of the FLUX project was to split completely the business content from the technology used for transporting and exchanging that content.

4.2. The FLUX TL conventions and behaviour

The FLUX approach is to create a business agnostic transportation layer facilitating the exchange of XML data between parties. In fact, this transportation layer could be used for any business (e.g. exchange data on airplanes).

The FLUX transportation layer basis provides description for:

• The FLUX Envelope, one single yet universal message format (XSD) that can encapsulate any business-specific message or structured data in a predictable way whatever the business system and associated data types and formats, using industry standard data representation techniques

• The FLUX Protocol, a mechanism describing how to reliably deliver the FLUX Envelopes to their destination reliably and without human intervention, leveraging state-of-the-art existing technologies (SOAP Web Services) in a sensible manner so as to as much as possible avoid interoperability issues between FLUX implementations based on different vendors' solutions.

4.3. FLUX TL Software

A party's system that needs to interact with a system of another party using FLUX obviously needs to adhere to FLUX conventions and behaviours. In other words it must implement the FLUX protocol.

One solution is to have such a system incorporate a tailor-made module that implements FLUX messaging built according to the technical FLUX specifications.

In many cases however, the cost of adding a custom-made FLUX module in an application can be very expensive, even impossible in case of legacy systems built a long time ago and for which the detailed technical expertise or knowledge of its internals is no longer available

Therefore DG MARE may provide a reference implementation of a FLUX Gateway that can on one hand communicate with legacy systems, and on the other hand connect to the FLUX TL.

A first Gateway (or bridge) is already available to exchange VMS data on the FLUX TL.

4.4. FLUX network

Parties can install this reference software, connect it to their local IT systems, and start communicating (using the FLUX protocol) with other connected parties. Such parties are called "nodes" or "endpoints".

One difficulty is the maintenance of such network.

As an example:

> In the EU alone, 24 countries are involved in fisheries and are exchanging data with each other, with DG Mare, with Eurostat, the EFCA, ICES and other organisations. There are about 13 RFMO with a large number of contracting parties and a large number of bilateral fisheries partnership agreements.

A typical EU country exchanges data with +/- 15 parties. This means that, working bilaterally, 22*15 connections, or 330, have to be made to cover for existing practices. Any changes in business could lead to a growing need for more connections.

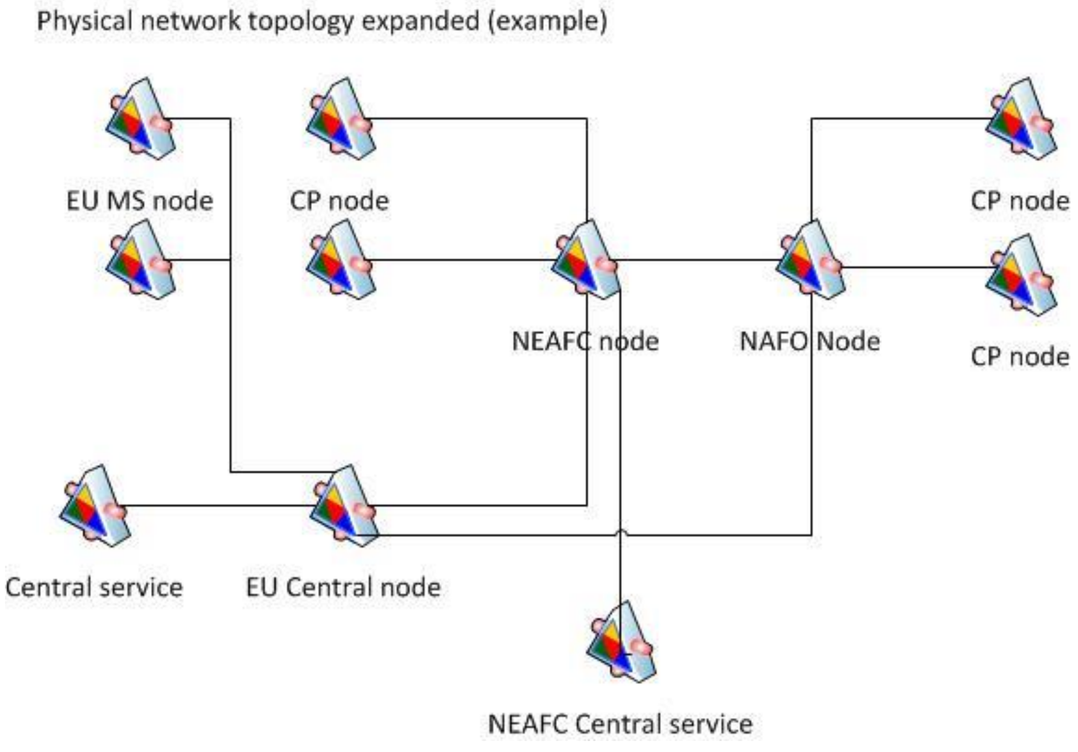
The cost of establishing (and maintaining) all these bilateral technical "links" between each parties' computer systems grows exponentially with the number of parties involved and prohibits setting up an efficient system.

To avoid this exponential growth an architecture is set up grouping parties around "central nodes". These central nodes can themselves connected to other central nodes. This means that if a party is connected to one central node, it can securely communicate with all parties connected to any central node. This approach is called the "data exchange highway".

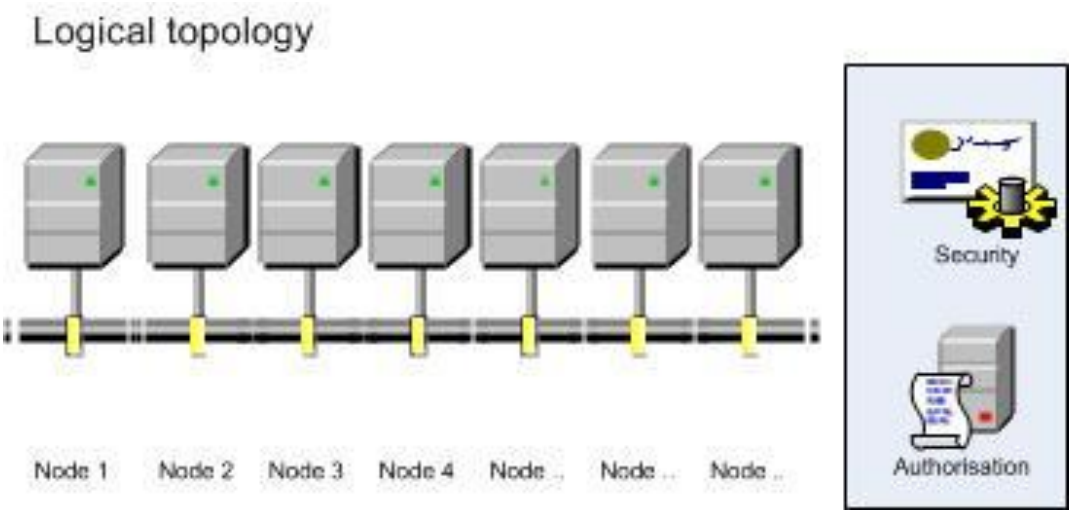
Typical candidates to provide such a central node are RFMO, or similar organisations.

4.4.1. Physical topology

The below drawing shows one possible example of such network. The software can be configured to make any physical topology from fully connected to a star configuration.



4.4.2. Logical topology



No matter how the physical topology is made, the software can be configured so that it appears that every node is directly communicating with any other node.

Hence, Physical topology becomes "unimportant"

- Configuration can be adapted to new needs
- Data flows can be optimized
- Network can organically grow

Each node can be configured according to its role

- Load balancing between instances
- Automated data flow optimisation

4.5. Conclusion

Parties having the transportation layer in place should never have to wonder again how they will exchange data for any data exchange requirement.

The software can be maintained and upgraded by IT specialists without hindering business people.

Data exchange protocols are identical for the whole business; the data itself can be freely defined for each particular case.

5. UN/CEFACT STANDARDISATION

5.1. The UN/CEFACT context

The United Nations Economic Commission for Europe (UNECE) is a focal point for electronic business standards, dealing with both commercial and government business processes to encourage growth in international trade and related services. In this context, UN/CEFACT (the United Nations Centre for Trade Facilitation and Electronic Business) was created.

UN/CEFACT is an intergovernmental body of the UNECE Committee on Trade, aiming to assist transactions through the simplification and harmonisation of processes and information flows. One of the working methods consists of developing methods to facilitate the transactions, including the relevant use of information technologies.

In order to secure coherence in the development of Standards, UN/CEFACT is cooperating with e.g. ISO (International Organisation for Standardisation) and IEC (International Electro technical Commission). UN/CEFACT has developed instruments such as XLM schemes providing a series of coherent, consistent and normalised syntax solutions that are aligned with domain reference models for publication.

5.2. UN/CEFACT standardisation for fisheries

UN/CEFACT standardisation perfectly suits the technology that was already selected earlier for the already existing ERS systems – with respect to the use of XML.

Even more important, the UN/CEFACT sectorial Agriculture group (under which fisheries standardisation resides) involves standardisation experts with relevant experience in fisheries and/or agriculture from all over the world and from different institutions.

Last but not least, UN/CEFACT standardisation is gaining traction from different environment like tax authorities or, food and health administrations. This makes the standardisation work for the internal needs of the fisheries community directly relevant in a much broader context.

5.3. The standardisation process

In a nutshell, the standardisation process has 3 main phases:

- The Business Requirement Specification
- The Harmonisation Process
- The Publication of the standard

The outcome of this process is a general toolbox of data elements and data exchange messages. An Implementation Document, made in parallel, describes how this toolbox is to be used within a given context.

5.3.1. The Business Requirement Specification (BRS)

During this phase, experts in the field (e.g. data exchange requirements for tropical Tuna) sit together and define the reporting needs for their field. In theory, they can do this without much knowledge of the UN/CEFACT standardisation process. However, knowing the UN/CEFACT philosophy, and having experience with the next phases, is hugely beneficial for making a BRS document fit for the purpose of standardisation.

UN/CEFACT is requiring the use of the Unified Modelling Language (UML) and has prepared a document template for this phase.

At the end of this phase there is a single document listing all the data exchanges that have to take place between the involved stakeholders in the field, and detailing each data element that has to be contained in those exchanges.

5.3.2. The Harmonisation Process

Harmonisation ensures that completely unrelated standardisation efforts still lead to compatible results.

UN/CEFACT maintains a library (currently approximately 14 000 entries) of data elements and components that are the outcome of earlier standardisation projects.

The main objective of the harmonisation phase is to investigate how the data elements, messages and even processes described in the BRS relate to the components available in this "Core Component Library".

Where possible, data elements in the BRS are replaced by these already existing core components. Where core components are missing, these are added to the next release of the Core Component Library. Each year, two versions of that library are released containing the latest updates.

The end result of this phase is that a Requirement Specification Mapping (RSM) is produced which integrates the business requirements with the UN/CEFACT Core Component Library.

The impact and importance of this operation should not be underestimated.

As one example:

> When writing the BRS for the VMS domain we described a data element called "vessel position" containing the longitude and latitude of the position of the vessel. During harmonisation, this data element was replaced by the already existing component "GeographicalPosition".
>
> We can now exchange any vessel position with any organisation using UN/CEFACT standardisation before even knowing who those organisations are. This does not bring any new obligation for data exchange but makes the systems that we will develop implementing he standard fairly future proof.

5.3.3. The Publication of the standard

The most important publication is an XSD (XML schema definition). This is a technical file that can be used directly to exchange data, or, more likely, from which more restricted XSD can be defined for practical data exchanges in the newly standardised field.

UN/CEFACT does NOT impose the technology to be used for exchanging data. Theoretically, one can imagine sending business content around using email, on CD using postal mail …. But the XSD is, of course, most suited to be used in combination with SOAP or REST web services (WSDL).

5.3.4. Standard expansion

It can happen that a new standardisation effort in a given field is very closely related to another field already standardised. In that case several approaches are feasible:

The minimalistic approach whereby the business requirement specification of the new field is written, completely ignoring the already existing BRS, RSM and published standard.

In that case, harmonisation will ensure that the new field will use the same core components as used in the existing standard. However, in this case there is limited harmonisation of business practices.

The maximalist approach whereby the existing BRS and RSM serve as starting point for the new standardisation effort.

In areas like fisheries, where many parties are operating on a global scale, the second approach is to be recommended as it facilitates re-use of software modules, reduces the need for training of staff and accelerates the standardisation process.

5.4. Conclusion

UN/CEFACT standardisation is a key feature for organising data management and data exchange on a global scale. It allows completely independent groups of stakeholders to build a global data exchange "language" for fisheries that is also compatible with related sectors where the same efforts are being made.

6. STATE OF PLAY

6.1. FLUX Transportation Layer

The root of the transportation layer software is the FIDES system, in operation in the EU since many years and intensely used for data exchange between the EU Commission and Member States.

A predecessor of the FLUX TL, called "Data Exchange Highway", is used for EU-ERS and NOR-ERS. On a yearly basis EU member States and DG Mare are exchanging 1 million messages on this network.

The FLUX TL will replace both above mentioned networks by the end of 2015 for all current data flows, while new data flows (not yet existing) will be added in the same period.

FLUX TL version 1.2 is currently being rolled out to EU Member States. On 27/03/2014 there are already 10 Member States exchanging data (ACDR) with DG Mare.

FLUX TL Version 1.3 (mainly bug fixes and performance enhancements) is scheduled to be realised by the end of June 2014.

6.2. UN/CEFACT standardisation

6.2.1. General principles

First, a number of general principles were defined to which each and any data exchange message has to comply. These general principles are the basis for all data exchange messages.

One interesting feature is that each data exchange message contains a GUID (Global Unique IDentifier). The GUID and the business content of the message may never be split, and in case that a received message is forwarded, it has to maintain the same GUID.

This ensures that it becomes very easy to discover that some data has been received multiple times, and it facilitates correcting earlier messages by simply mentioning the GUID of the previous message.

Most modern IT languages contain functions to generate GUID.

6.2.2. Standards

UN/CEFACT standards are already available for VMS, MDR and ACDR (aggregated catch data reports);

Additional standards are expected to be ready for fishing activities, landing declarations and sales notes by the end of 2014.

One project, FLEET, is on-going and in cooperation with FAO (Global Fleet Register initiative), a UN/CEFACT standard is created for the exchange of vessel data.

The LICENCE project aims at creating a complete electronic workflow over the FLUX TL (and using UN/CEFACT standardised messages) for fishing licences and authorisations.

Several working groups are investigating other domains like e.g. the data exchange requirements for Bleu Fin Tuna (possibly to expand to cover also tropical Tuna),

6.3. Central web services

The aim is to make a number of central services available on the FLUX TL. These could range from Geographical Information Services to advanced calculators for scientific purposes depending on the needs that arise. For now, first experiments are made with two services:

- MDR: Master Data Register containing all the code lists that are needed for all data exchanges over the FLUX TL. This service is available on the FLUX Transportation Layer but still needs to be officially released. This is foreseen for the near future.

  The MDR contains all code lists to be used by EU vessels. The existence of this MDR does not prohibit another organisation to bring another MDR on-line for its own needs.

- FLEET WS: This is a temporary web service delivering data on vessels from the European Fleet register. It will be replaced by the outcome of the FLEET project later.

6.4. Implementation

The short term implementation Plan for the EU looks as follows:

Data Flow	Testing		Production	
	Start date	End date	Start date	End Date
ACDR	10/01/2014	15/02/2014	17/02/2014	17/02/2014
MDR	1/05/2014	30/06/2014	5/05/2014	30/08/2014
Fleet WS	1/05/2014	30/06/2014	5/05/2014	30/08/2014
VMS	1/05/2014	30/06/2014	15/06/2014	30/11/2014

7. CONSIDERATIONS FOR THE FUTURE

7.1. FLUX Technology stack

The FLUX TL reference implementation is based on an open source technology stack. All contributing parties are invited to use the same technology stack for their local developments directly related to the use of the UN/CEFACT standardised messages.

7.2. Modular framework

Not every party is involved in all aspects of the fisheries business. As one example, a country not involved in Bluefin Tuna fishing has no interest in implementing the data elements, messages and processes for that particular fishery.

The FLUX business layer is based on individual stand-alone business modules that allow parties to implement only the modules they need. The UN/CEFACT standardisation approach guarantees that modules are compatible.

Once a party has completed a FLUX data exchange installation for a single module, it should be fairly easy to plug in extra modules for other data exchanges.

The intention is to work together on building modules that can be integrated with the FLUX TL reference installation. Such modules could be Electronic logbook or VMS viewers, Inspection report management software, licence management tools, vessel management software …

7.3. Open source community

Finally, once the first modules being sufficiently stable, the intention is to construct an open source community around the software involving, one way or another, the cooperating parties. The modalities of such community have not been considered to any detail yet but the European Union Public Licence is obviously a good candidate licencing scheme.

8. ADDITIONAL INFORMATION

The UN/CEFACT website

http://www.unece.org/cefact.html

The Master Data Register static pages

https://circabc.europa.eu/w/browse/3cc8c417-0f2a-4eb4-8ff7-10d60638446a

Information on the EUPL

http://opensource.org/licenses/EUPL-1.1

9. ANNEXES

9.1. Basic technology

9.1.1. XML

FLUX is based on the use of XML. XML stands for eXtensible Markup Language and is a textual format which was originally designed to describe, transport and store data. The basic syntax is very simple. As soon as two parties agree on a set of "tags" they can exchange data.

As a basic example, we could have the following agreement to describe a person's data:

```
<person>
   <name></name>
   <email></email>
   <tel></tel>
</person>
```

Starting from the above XML Template it becomes easy to exchange data. It only takes entering data in between the tags.

```
<person>
   <name>firstname lastname</name>
   <email>name@somewhere.eu</email>
   <tel>+xx yyy zz tt uu</tel>
</person>
```

The advantages are immediately visible:

The encoded data is easily readable by a human being (in contrary to other formats)

One can send an xml file over the internet between completely different computer systems. In fact, once an XML template has been agreed, computer systems which are otherwise not compatible can still exchange data provided that both systems contain a "parser" [a small computer programme able to break up the xml file into its individual components and feed then to the local computer system, or vice versa, combine local data in an xml file to be sent to the other party.]

The flexibility offered by XML has made it very popular and many computer languages contain ready-made tools facilitating the use of XML. In fact, the most current versions of office suites (like Microsoft Office and Openoffice) store data and texts in a native xml based format.

9.2. Standardisation

The above example shows the huge flexibility of XML. It also demonstrates one of its biggest flaws. To be able to exchange data between computer systems, the sender and the recipient have to agree on a set of tags and on the attributes and characteristics of the data that may be contained within those tags. For data communication between many parties, often spread around the globe, this is virtually impossible.

Moreover, it was soon discovered that XML could be used for other purposes than exchanging data. For instance, a company could use it to document services it was offering on the internet or to explain to third parties how its database was structured.

The World Wide Web consortium (W3C), the internet standardisation organisation, soon developed a number of standards in various XML related fields. Within the FLUX context the following are the most relevant:

XSD (XML Schema document): an XML based language to describe advanced data structures and the validation rules applicable to them

WSDL (Web Services Description Language): an XML based language to describe web services.

The XSD describes which messages we are sending around (COX (catch on exit), COE (catch on entry), DCA (daily catch activity), POR (port)…) the data included in those messages and the validations in place.

The WSDL describes which functions are available to process sending and receiving the data contained in the XSD. Or, the WSDL is describing the behaviour of the computer systems used to transmit data.

Web services are programs available on a computer connected to the internet. These programs can be used by any computer connected to the net provided that this second computer is given the necessary authorisations and technically knows how to "talk" to the other computer.

These computer "conversations" are described in "protocols". One of these protocols is the SOAP protocol which is used for FLUX. It could be seen as a kind of "grammar" for computer talk similar to what grammar is for English or French.

9.2.1. Specialised software

This standardisation, and the popularity of XML in general, has motivated a number of companies, and open source development teams, to make specialised software packages facilitating the use of XML.

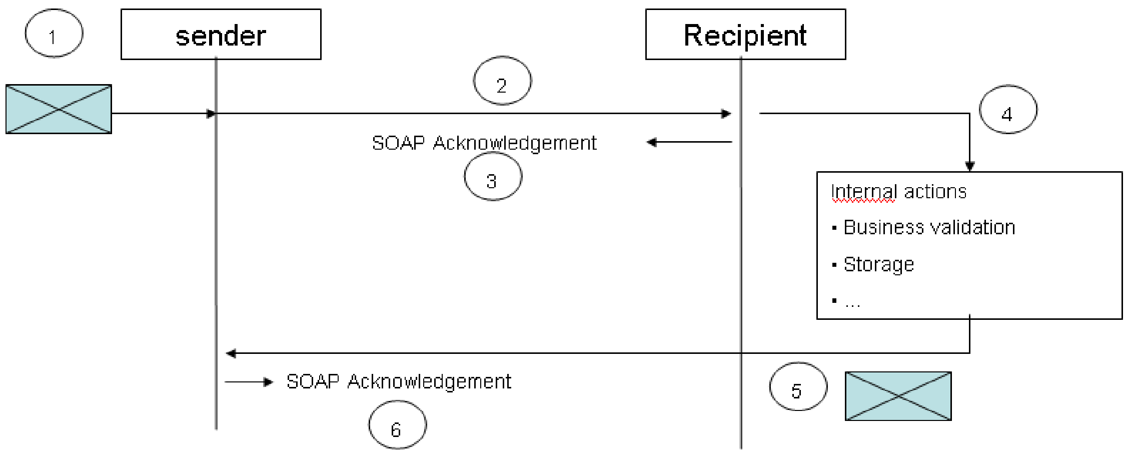
This means that, once WSDL and XSD available, the individual party does not have to develop a complete software solution from scratch, but can start from often already very advanced standard software packages able to "understand" a WSDL/XSD and "deploy" the needed web services (including error handling, rejection of faulty messages…).

The main task of the individual party is to make a connection between these packages and their local systems, and to ensure that the right data is included in the XSD templates.

However, the importance of these packages should not be overstated and there is a substantial learning curve involved in configuring the data exchanges. The experience with FLUX, made available to the fresh starters, will be very beneficial to a smooth introduction of these technologies.

9.2.2. *A typical data transfer*

A data transfer happens typically between two computer systems connected to the internet which have both been configured with the same WSDL and XSD. We distinguish in the drawing below between the "sender" and the "receiver".



The sender prepares a message with content compatible with a message description in the XSD

… and sends it over the internet to the web service of the recipient.

The recipient's system checks whether the message is complete, and whether it responds to the validation rules included in the XSD, and sends a SOAP acknowledgement. This typically takes a few seconds. This SOAP acknowledgement can be compared to the "delivery" message of email systems; It does NOT mean that the recipient agrees with the content of the message but only states that the message was technically correct and well-received.

The recipient goes through a series of internal business processes. Dependent on the case this can be fully automated, or require human intervention, e.g. giving a prior authorisation to enter the waters of another party.

The recipient prepares and sends a Return message containing the business evaluation and reaction to the sender's message; this message can be very simple saying "OK" or can be very complex sending a large amount of data back to the sender.

The sender sends a SOAP Acknowledgement to the recipient confirming that he has well-received the Return message.

The SOAP acknowledgements are called synchronous responses because these happen typically within seconds after receiving a message and within the same "session" of the SOAP protocol.

The Return message is called an asynchronous message as it can be sent after a fairly long time (ultimately days). It is of importance to agree upon the maximum delays for sending a Return message.

These Return messages are needed to allow for a more thorough, and time consuming, evaluation of received messages than what is normally acceptable within a SOAP session.

In general, we distinguish two forms of communication:

The "push" mechanism meaning that the data creator/owner is sending (pushing) new data to the receiver(s) needing that data. In this case the content is contained in the first message while the return message is simply saying "ok", or gives a number of error codes in case that the data sent contains errors.

The "pull" mechanism is used for situations where the sender is actually asking (pulling) the receiver for a dataset available on the receiver's computer system. In this case the first message is only a simple request, replied to with a possibly very large and complex return message containing a data set.