



## Appendix 1: EMREX Description

### Description of ordering organization

CSC - IT Center for Science Ltd. is a non-profit, state-owned company administered by the Ministry of Education and Culture. CSC maintains and develops the state-owned centralised IT infrastructure and uses it to provide nationwide IT services for research, libraries, archives, museums and culture as well as information, education and research management.

CSC has the task of promoting the operational framework of Finnish research, education, culture and administration. As a non-profit, government organisation, it is our duty to foster exemplary transparency, honesty and responsibility. Trust is the foundation of CSC's success. Customers, suppliers, owners and personnel alike must feel certain that we will fulfil our commitments and promises in an ethically sustainable manner.

CSC has offices in Espoo's Keilaniemi and in the Renforsin Ranta business park in Kajaani.

More information at [www.csc.fi](http://www.csc.fi)

### Overview of Emrex

EMREX is the solution for electronic transfer of student records between higher education institutions in Europe. The EMREX project addresses the EU 2020 target that 20% of higher education students should be mobile. The biggest benefit of EMREX will be the increased availability, quality and reliability of information about student records of achievement information.

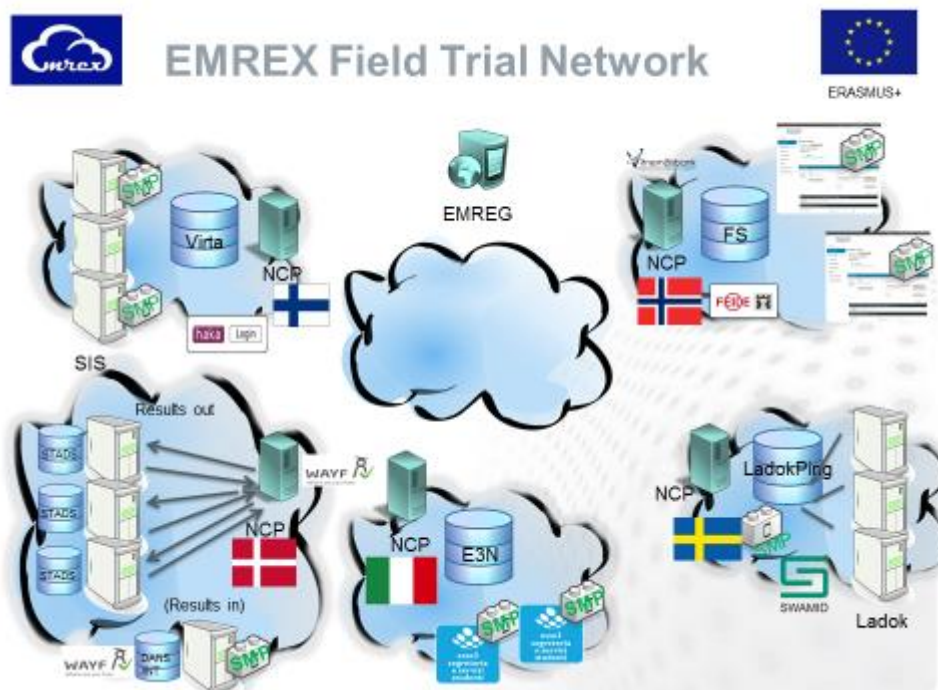
The EMREX field trial aims at testing new ways to make the administration of student mobility easier and thus promoting higher attainment level to student mobility in higher education and also encouraging more effective recognition of prior learning and avoiding overlapping studies. In the first phase the transfer will be set-up between four Nordic countries and Italy.

More information to be found from the project workpages, [www.csc.fi/emrex](http://www.csc.fi/emrex)

## System description

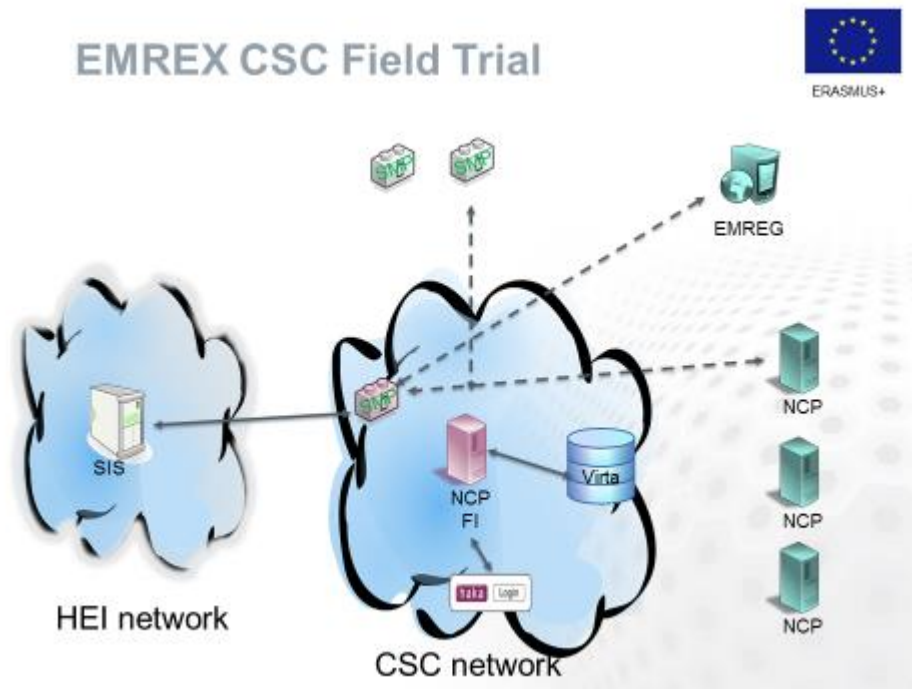
The EMREX field trial will be conducted between four Nordic countries and Italy. The field trial will be evaluated by the University of Warsaw in Poland. Each participating country will develop its own part of the system, although the aim is to re-use much of the work in other countries.

Below is a picture of the field trial network and the nodes involved. The setup might change as more information becomes available.



Picture: EMREX field trial network

CSC is responsible for the development and implementation of the components used in Finland. That is, the Finnish NCP and the SMPs used in Finland. The SMPs will be connected to the Higher Education Institutions' (HEIs) Student Information Systems (SIS). The following picture describes the solution from a Finnish perspective.



Picture: CSC Field trial implementation

## Current situation

Currently study records (achievement records) are not transmitted electronically between countries. Instead paper copies are used. In Finland the study records can be found centrally, from the National Data Warehouse for Higher Education (a.k.a Virta). There is also a federated identity solution in Finland used by the HEIs, called Haka. The EMREX solution will utilize both these existing solutions.

## National Data Warehouse for Higher Education (Virta)

VIRTA is a centralized database in which institutes of higher education (HEI) keep partial copies of their student and other person registries. Altogether, achievements of 38 HEI's in Finland, incl. 1.3M students are included in VIRTA.

The data is securely and efficiently available for agreed third parties via a technical user connection. VIRTA provides a method to organize data harvesting in the way which improves quality of information and decreases workload of the universities.

CSC- IT Center for Science Ltd. is held responsible for developing and maintaining the Finnish national data warehouse for higher education. Access to VIRTA database is controlled by the Ministry of Education and Culture in Finland.

Each institute of higher education is responsible for collecting the data in a commensurable format as well as delivering the information to the National data warehouse for higher education. Target for data update interval is one day.

As a project outcome, the National data warehouse for higher education exists including a joint data- architecture, data integrations by each higher education institution, a fully centralized data warehouse and data flows by the authorities agreed to be conveyed from the data warehouse. Joint Data-Architecture is part of a so called “XDW” conceptual formula.

Parallel to the project a national individualization service, a so called national student identification (OID) will be taken into use for students with no social security ID (i.e. exchange students)

More information:

<https://confluence.csc.fi/display/VIRTA/National+Data+Warehouse+for+Higher+Education>

## Federated identity solution (Haka)

Haka is the identity federation of the Finnish universities, polytechnics and research institutions. Users are able to access federation services using a single user account and password. User identities are provided by the users home organizations

Haka federation is built on [SAML](#). Majority of Haka members are utilizing [Shibboleth](#), a standards-based open source middleware software. Haka federation operator CSC maintains information on federation members and partners. This information is published in Haka metadata. End user attributes applied in Haka are defined in FunetEduPerson schema.

Haka SAML 2.0 profile is based on the common SAML 2.0 profile of the Finnish public sector

More information from the Haka wikispaces,

<https://confluence.csc.fi/pages/viewpage.action?pageId=29395721>

Haka federation is operated by CSC - IT Center for Science Ltd.

## HEI SIS to connect to

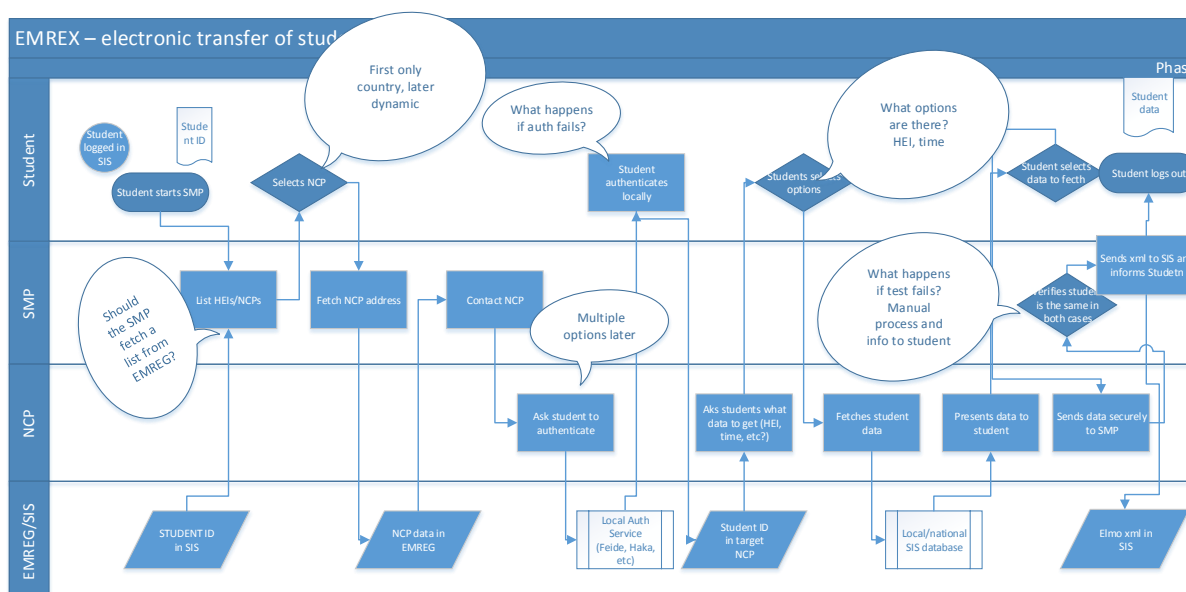
The EMREX solution will need to communicate with the HEI SIS systems. At the moment the following SIS are considering participating in the field trial

- Oodi (Hanken)
- Asio (OAMK, JAMK, KAMK, MAMK)
- Opsu (TaY)

- Oprek (TTY)
- Sture (Åbo Akademi)

## EMREX process

The EMREX process for transmitting study records between HEIs is described in the flowchart below.



Picture: Emrex flowchart

## The key user stories

Recognition of previous studies for an active student (main focus in EMREX)

Geir is an active (enrolled) student at the University of Oslo in Norway. As part of his studies he spent an exchange period in Finland, where he took courses from Helsinki University and the Metropolitan University of Applied Sciences. Now Geir has returned to Oslo and would like to get his studies in Finland recognized as part of his degree program in Oslo.

Geir logs onto StudentWeb (connected to FS - the SIS system of the University of Oslo) using his local credentials. He then starts the Student Mobility Plugin (SMP) by clicking the correct link in StudentWeb. Thus Geir is already authenticated in SMP with his local credentials.

The SMP asks Geir to choose from which country (note: in future releases there could be a dynamic selection process, where the students selects the university, region and/or country



and the system then selects the correct NCP) he would like to retrieve his student data. Geir chooses Finland.

The SMP contacts the EMREG register to find the address to the correct National Contact Point (NCP). After that the SMP contacts the NCP to inform the NCP that a student called Geir is asking for his student data in Finland.

The NCP asks Geir to authenticate himself in Finland. Geir selects Haka, the Finnish authentication service, (in the trial only Haka will be supported, later perhaps other authentication services) and enters his username and password that he used, when studying in Finland at Helsinki University. (note: Geir needs to do this before his account in Finland expires). The NCP confirms that Geir is authenticated in Finland. The NCP then asks Geir what information he would like to retrieve. (note: still to be decided: suggestion is to possibly first choose from which HEI and then filter achievements based on HEI, date, degree). Geir selects the courses he wants to get.

Using the Finnish credentials the NCP then contacts the Finnish national data warehouse, Virta, and asks for Geir's data. The received data is then presented to Geir. The NCP also checks whether course descriptions are available and adds that to the data when possible (NCP specific).

Geir confirms the course data he wants to have.

The NCP signs the student data and sends it back to the SMP in a secure way.

The SMP performs a check based on name, birthday and sex to confirm that the Geir in Finland is reasonably the same as the Geir in Norway. If this check would fail, the SMP would flag the data and an administrator would have to manually confirm that the two Geirs are the same person.

However, as the check is successful, the SMP then saves the student data in the correct local database in FS. The SMP informs Geir that the retrieval process was successful and that the data is now in the Norwegian recognition process for approval by an administrator. If appropriate, the NCP also sends a trigger to the local SIS that there is new student data for processing.

Geir exits the SMP and continues the process in StudentWeb.

Student admission: eligibility and/or reviewing student data (similar EMREX-process, but could be started form another system)

Pamela has a BA from the University of Gothenburg. She would like to continue her Masters studies abroad and is applying to the University of Copenhagen.

Pamela starts the application process using the admission system in Denmark. When filling in the application form, Pamela is asked to present proof of her BA degree. Pamela then clicks a link in the admission system to electronically transfer her degree data. The link opens up the



same SMP webpage as in the first case and from there on the process is similar to the first case.

Post-studies retrieving student data (optional, not implemented in EMREG)

Simone is applying for an international job in marketing. The company has operations in Poland and Simone would like to prove that he has indeed studied international cuisine at the University of Warsaw.

Simone opens a national SMP-service page. He is then asked for what purpose he is retrieving his study records and selects that he wants to show his record to a third party.

Simone is then asked to authenticate himself using a national authentication service. After identification he selects the country from which to retrieve his information. Simone selects Poland.

From here on the process is similar to the first case. Simone authenticates in Poland (note: requires lifelong access rights, or that there is a EU-wide eID in place). He can then select what data he wants to retrieve and the data is sent back to the national SMP service.

The national SMP service then stores the data securely and presents Simone with a corresponding access code to the data. Simone can then give this access code to his employer.

The employer also goes to the national SMP-service page and enters the access code. The employer is then presented with the verified achievement records of Simone.

After a set date the access code expires.

## Use Cases

		<b>Active student fetches study records from another university to get the previous studies recognized in current program</b>	<b>Logged into SIS</b>	<b>Receives Student Records as XML or PDF</b>	<b>1 Has access rights to needed data and user name in SIS</b>
1	Student	Student starts SMP in University	Student logged into University SIS	SMP started	1 For Emrex field trial it is assumed that only a SMP connected to a SIS in used so that the ID can be transferred
2	SMP	SMP receives student ID from SIS	SMP started	Active session in SMP with Student ID info	1 In future releases user might need to authenticate in SMP if no previous auth is done
3	SMP	SMP fetches list of available NCPs/countries from EMREG	EMREG up-to-date	List of available countries and corresponding	1 EMREG can return a list of available countries

				NCPs with details	
4	SMP	SMP presents student with list of available countries	List of available countries		1
5a	Student SMP	Student selects country to fetch records from		NCP selected	1 Note: in Italy the student will probably need to select HEI as well
5b	Student SMP	Optional: Students selects HEI/region and/or country and SMP determines the corresponding NCP		NCP selected	3 A more dynamic selection process for future development
6	SMP	SMP contacts correct NCP	NCP selected and IP-address and public key available		1
7	NCP	NCP presents student with authentication options			1 In beginning only one option, but in future releases there could be more
8	Student NCP	Student authenticates in NCP. NCP contacts auth. service	Correct NCP contacted	Student authenticated	1 Student must still have access rights in that country (NCP)
9	NCP Student	NCP asks student for local selection criteria (e.g. HEI) before fetching results. Student makes selection		Search criteria selected	2 Optional when locally needed
10	NCP Database	NCP fetches student's records from database	Student authenticated in NCP, selection done	Records in NCP	1
11	NCP	NCP presents the available records to the student			2 If locally implemented
12	Student	Student filters records and selects which study records to import		Study records selected for transfer	1 Filtering done by HEI, date and degree
13	NCP	NCP attaches course descriptions to selected courses, if available		Course descriptions attached to ELMO	2 When possible, the NCP should try to contact correct parties to fetch course descriptions. Probably correct HEI directly or national database.
14	NCP	NCP signs and sends selected courses to SMP as XML			1 Signed and using secure transfer
15a	SMP	SMP checks that the selected student is the same as the one in			1 Note: still to verify how the matching process should be done, especially if it fails



		SMP and result is positive			
15b	SMP	SMP check fails. Process stops and student is informed that manual intervention from admin is needed	Check fails	Process halted and information package presented to student	Suggestion only at this stage: If check fails no data can be transferred to SIS. The process stops and the student is informed. The student can then ask for manual matching by HEI's admin. If manual matching is ok, then admin can restart the process
15c	Admin	Admin manually approves transfer in SMP	Check has failed	Process continues	Suggestion still. Other possibility is that the data is transferred to the HEI and an admin get it from there if the manual check is successful
16	SMP	SMP presents student with imported data, including course descriptions			2
17	Student SMP	Student checks that the received student records are the ones he wanted			2
18	Student SMP	Student import the data from the SMP to the local SIS		Study records in SIS	1 Could be data import or saving xml.pdf to a certain folder
19	Student SMP	Student exits SMP		Process ended	1 Student can also exit SMP without importing data to SIS
20	Admin	Add/delete admin user rights in SMP, including who can manually restart halted process			2 Depending on how the SMP is implemented. Could be part of the SIS admin also
		<b>Administrator maintains NCP list in EMREG</b>			
1	Admin	Verifies new NCP and adds to list			1
2	Admin	Changes information on existing NCP			1
3	Admin	Deletes NCP from List			1
4	Admin	Add/delete admin user rights			1
		<b>Administrator maintains NCP</b>			
1	Admin	Add/delete admin user rights			1

2	Admin	Retrieves log for a specific student getting his student data	1
3	Admin	Add/changes local filtering criteria	3

## System Architecture

The latest technical descriptions and planning can be found here: <https://confluence.csc.fi/display/EMREX/Planning>

The EMREX solution will be developed using open source technologies. The NCP in Finland will be done with Java. The SMP in Finland will be developed using java and/or javascripts.

Simultaneously with the development in Finland similar components will be developed in Norway and possibly in Italy, Sweden and Denmark. The goal is to reuse as much as possible of the development in other countries and vice versa.

The preliminary data flow is shown in the figure below.

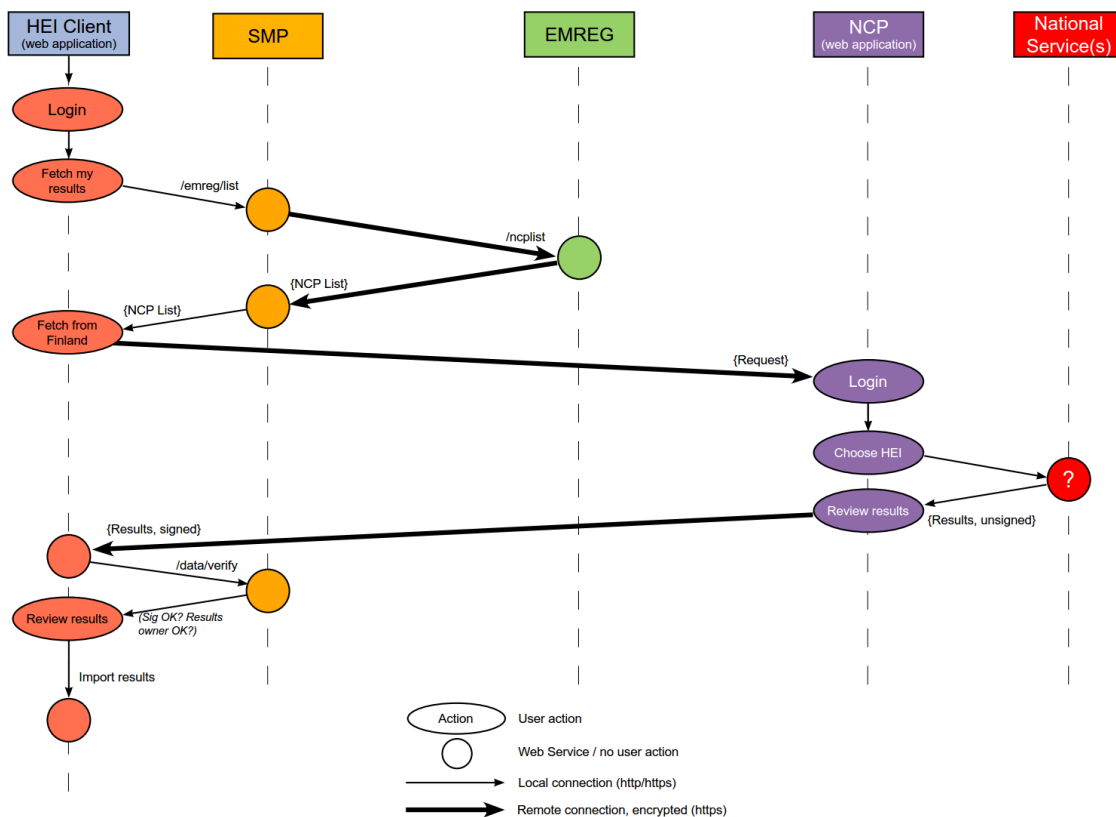


Figure: Data flow

The system architecture of the NCP and SMP components to be developed are shown below.

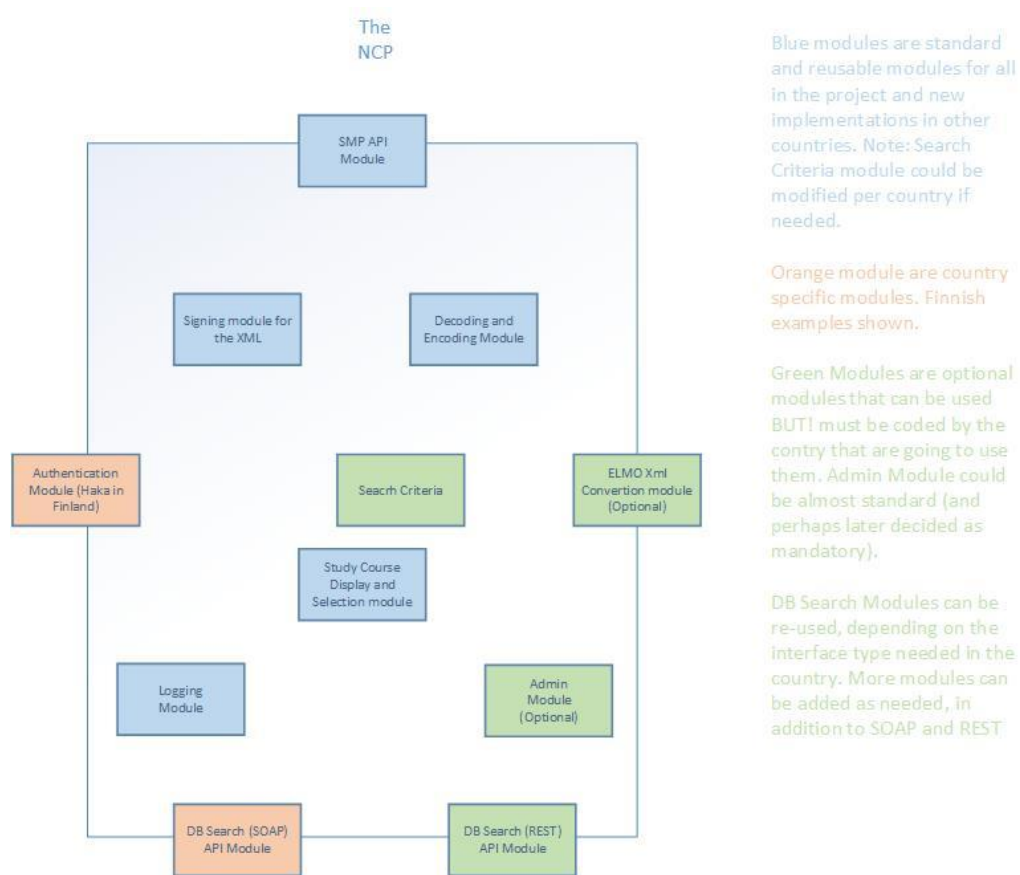


Figure: NCP architecture

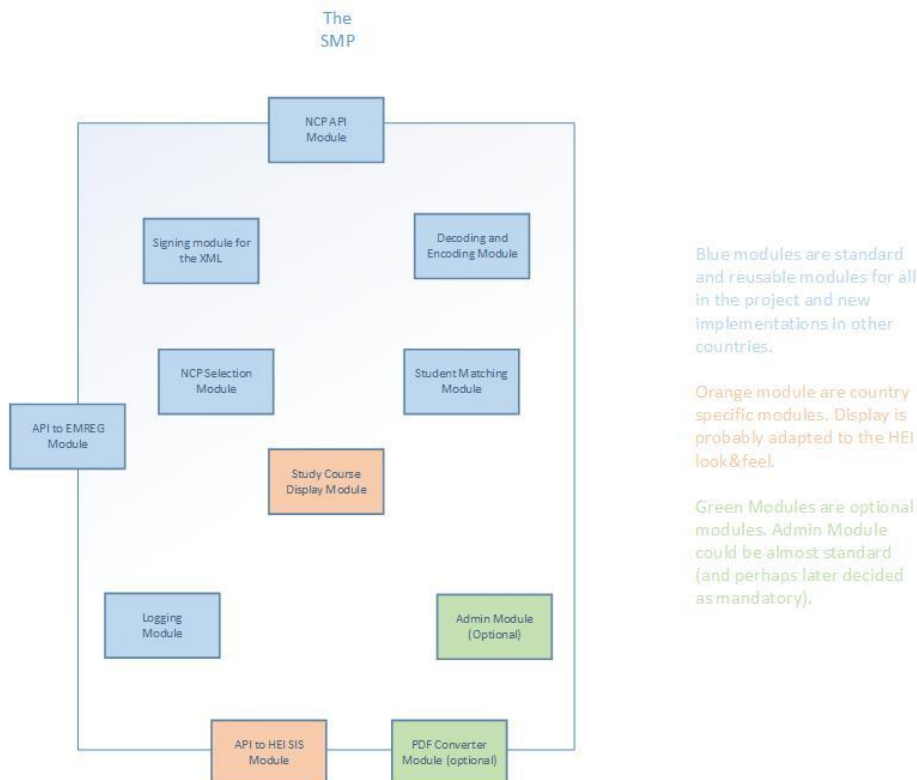


Figure: SMP architecture

## System

- The system is modular (SOA), see system architecture pictures
- Each module must be done in a way that makes it possible to re-use it in new implementations (assuming of course that the next implementation also uses e.g. java)
- The system shall support the user stories and use cases described in the planning section

## Capacity

- For the field trial only a few hundred users are foreseen, and only a handful simultaneous users. This number is expected to grow after full production is reached. Especially the NCP needs to be able to handle hundreds of sessions. Also the EMREG needs to be able to handle a large number of SMP requests.
- EMREX is a real-time system. Response times shall follow good standards for web applications
  - If fetching of some data is slow, the user should be informed with some sort of "waiting warning" (e.g. a nice movie)
- For the field trial only minor storage needs are foreseen, i.e. logging and possibly failed matches
- The system shall be scalable, so that capacity can be increased when needed
  - For the field trial the number of simultaneous users will be very low (less than 10?)

- For future use, peak time could be when there is a deadline for admission
- EMREG, NCP, SMP have different capacity needs
- Availability. EMREX is not a critical system, normal uptimes for web applications are sufficient. However, in the future it is possible that there will be time slots when the system must be functioning, i.e. during peak admission application periods.

### Security

- Enforce https on all SMPs and NCPs, a requirement to be able to use EMREX at all
- The only crypto stuff we really need to do is to sign the ELMO document using xml-dsig; for archival purposes, the document should include a timestamp and certificate as well; having the signature be part of the document itself is a good solution
- A centralized list of trusted NCPs and their certificates (that is, EMREG)
- In addition there can be local security rules that needs to be followed to be able to connect Emrex to existing systems
- The student is authenticated both in the SMP (home country) and NCP (or host country HEI)

### Logging

- For the field trial we will do local logs in each component. Logs will be stored at least for the duration of the field trial
- Later we need to decide how long the logs needs to be stored (years?), or how long it is allowed to store them
- Logs will be in English, except for where local language is used (e.g. name of HEI)
- Before the project is finished we need to suggest how logs should be handled and also if central logging could be used and for what
- Logging also needs to take into account the evaluation part, i.e. need to log how many students used Emrex and from where (at least)

### NCP Logging:

- From where (school) the connection is done.
- Who (person) is trying to connect or connects.
- What is fetched.
- Date and time of the connection.
- From the Logs we have to count the number of connections per person per day/week/month.
- Authentication logging successful versus unsuccessful (Finland specific logging into HAKA).
- Failed logins the connection string is wrong or the authentication key is wrong.
- Matching of the person fails.
- No data is found in the student registry system (Finland VIRTIA).

### SMP Logging:

Scenario 1, SMP connected to the Schools system:

- What is fetched from the EMREG server.
- Who (person) is trying to connect or connects.
- Date and time of the connection.
- Failed fetches using wrong EMREG server-data.

Scenario 2, SMP is a freestanding multiuser plug-in:

- From where (school) the connection is done.
- Who (person) is trying to connect or connects.
- What is fetched from the EMREG server.
- Failed fetches using wrong EMREG server-data.
- Date and time of the connection.

### Testing

- First solution testing in test environment with test data
- NCP tested against another SMP (e.g. in Norway)
- Test network (EMREG, NCP, SMP) for testing
- Second pilot testing in production environment (or test environment in those cases that is contains real data)
- Test report, which includes test cases and results
- Capacity testing, penetration testing, failure recovery

### Other requirements

- Documentations
- The language used for documentation is English
- The final code should be stored in GitHub
- The final common code should be licenced under EUPL (i.e. as open source also afterwards)

### Requirements for Finland

- Java
- SOAP towards Virta
- Haka integration
- Eden development environment
- Platform/OS, network connections, storage





ERASMUS +



## Description of environment

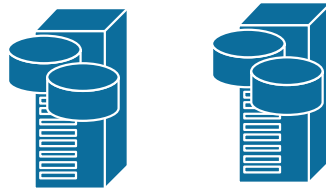
### CSC:n Development and Production environment

The service shall be developed using the CSC Eden development environment and then deployed in the CSC production environment.

For more information about Eden, see [ww.csc.fi/eden](http://ww.csc.fi/eden)



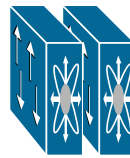
Redundant storage environment



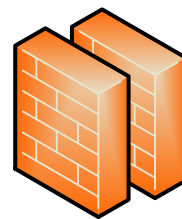
Redundant virtualisation environment



Redundant switching infrastructure



Redundant network load balancing



Redundant Firewalls (CSC)

Redundant routing with 10G connections

