

MLIS-5016 TransAccount

WP 5

Progress Report No 3

01.01.2001 – 30.06.2001

Version 1.0 from 06/08/2001

TransAccount Consortium

MLIS PROGRAMME

TITLE:	Progress report No 3, 01.01.2001 – 30.06.2001
PROJECT:	TransAccount MLIS-5016
AUTHOR:	TransAccount consortium
ESTABLISHMENT:	AT
ISSUE DATE:	06/08/2001
DOCUMENT ID:	TransAccount_05/01
VERSION:	1.0
DELIVERABLE:	(WP 5)
STATUS:	initial
DISTRIBUTION:	TransAccount, CEC
LOCATION:	ManagementReport3.DOC
KEYWORDS:	TransAccount, progress report

Date:	Version:	Author:	Comments:
06/08/2001	1.0	TransAccount Consortium	

**TRANSACCOUNT PROJECT
MLIS - 5016
MANAGEMENT REPORT No 3, as of June 30, 2001**

Content

I - CHANGES RELATING TO TRANSACCOUNT'S PARTNERS	4
II - TRANSACCOUNT'S PARTNERS ACTIVITIES AS AT June 30, 2001	4
EU (EUROSOURCES SA).....	4
EP (Alpnet Technology GmbH).....	4
SY (SYSTRAN SA)	5
III - TRANSACCOUNT'S PROJECT PROGRESS AS AT June 30, 2000	6
1 - Deliverables produced	6
2 - Deliverables to be produced	6
3 - Information and dissemination	7
4 - Cost Statements	7

I - CHANGES RELATING TO TRANSACCOUNT'S PARTNERS

EU – Eurosources / AM - Amyot Exco: Both partners have left the project and are in the process of being replaced by PricewaterhouseCoopers.

II - TRANSACCOUNT'S PARTNERS ACTIVITIES AS AT June 30, 2001

This section describes the activities performed by the partners in TransAccount project, for the period from January 1, 2001 to June 31, 2001. Under each partner's activities, we have mentioned the WP(s) in which those activities have been performed together with the corresponding deliverables.

EU (EUROSOURCES SA) WP1 - WP5

Eurosources is not any more partner in the project. No information is available therefore for this period.

EP (Alpnet Technology GmbH) WP 1 - WP5

- a) Update of the web server for TransAccount (www.transaccount.org)
- b) Replacement of the Unix based Globelix server by an NT version.
- c) Further porting and adaptation of Globelix server kernel to Windows NT to allow maximum access. A form based parser for accounts has been implemented. It is based on a formal description of an account in a form parser file. This file defines the various sections of the account and describes how to transform the forms in there (numbers) to another accounting scheme. The parser implemented reads the form parser file, in a next step applied the patterns given there to an instance of an account and creates an instance specific Perl script form that. This Perl scrip is then executed and creates the transformed account. As a next step the transformed account can be fed into the thesaurus based translator – applying TM and terminology functionality to the file. A prototype has been created using several GAAP accounts. Based on some sample data available from a Eurosources Geneter formatted terminology file the translation functionality could be successfully demonstrated.
- d) Adaptation of the TransAccount specific GUI based on Globelix. New features like access to the form parser have been implemented.
- e) Evaluation of the specifications of the other partners
- f) Alpnet has taken over the coordination of the project from Eurosources.

SY (SYSTRAN SA) WP3 - WP4

Project adaptation

Due to the withdrawal of Amyot, and difficulties encountered with the project management, and no delivery of the main resources needed to perform the specialised terminology insertion, text corpora and ontological modules, Systran had to look after a spare solution.

As it seemed that US GAP was not an actual standard, we tried to get some information about the International Accounting Standard that was being developed in Europe. Our discussions with PricewaterhouseCoopers about these topics resulted in a change in the project definition, the writing of an alternative technical definition, and the participation of this company to take over the remaining work to be performed by both Amyot and Eurosources.

Preparation of XBRL parsing

To anticipate translation of XBRL forms, we have adapted and developed a XML-based parser for XBRL handling (based on current specification draft: whatever will be the final specification for XBRL, different kinds of mechanisms have to be implemented). This parser converts the relevant part for automatic translation in the XBRL document into a raw textual document with internal Systran tags. The conversion does not include only content of entities, but also includes to attribute values needing to be translated. The insertion of internal tags guarantees that the re-introduction of the translated text into the original XBRL document is possible.

Moreover, this technical support of XBRL free text translation includes the handle of special XBRL entities such as period, unit, or currency : indeed, the translation of these tags is not relevant since the XBRL reader will handle them and expand them to natural language, but these tags are important to obtain a correct analysis of the sentences.

Example, for the following sentence, we have the following sequence: *During*
<period>xxx</period>, the society has...

Filter: *During* \$\$#period, the society has...

Translation: *Au cours de* \$\$#period, la société a...

Reconversion: *Au cours de* <period>xxx</period>, la société a

Internally, the \$\$#period reflects the presence of a <period> tag. Remark that the translation of the associate preposition is complex since, it really depends on the type of the <period> element (if it is a single day, a full year, or few months, the associated preposition won't be the same even if it is the same in source language. This axis for translation tuning is still under study).

III - TRANSACCOUNT'S PROJECT PROGRESS AS AT June 30, 2000

1 - Deliverables produced

The following deliverables have been due in this period.

Month 12:

- (a) Accounting Text Database (multilingual and aligned), datafile. This deliverable was in the hands of Amyot / Eurosources. No data file has been delivered so far from those partners. See also the comment at the beginning of the document.
- (b) Form based parser/generator for accounts for US and FR accounting systems, software. A report has been produced about the system architecture. The porting of relevant features of the Globelix software as the underlying software system has started and makes good progress. A version of the software was demonstrated at the review in June 2001.

2 - Deliverables to be produced

The www.transaccount.org web site is available containing all the deliverables existing. This is part of the management activities and updated by Alpnet. The site can be accessed by entering user name and password sent to all parties.

The following table shows the deliverables available for this period.

No	Specification Document	Due to in Months	Responsible	Deliverable Reference	Comments
WP 1/4	Multilingual terminological database	14	Eurosources	D1c	Geneter based data file
WP 2/8	Form Based Parser	12	Alpnet	D2a	Version 1.2; adapted version
WP 2/10	Form Based Translator - User Interface Manual	18	Alpnet	D2f	Version 1.3

3 - Information and dissemination

No special actions have been done in this period

4 - Cost Statements

The cost statements corresponding to the works and progress are attached to the present management report. They reflect the work done and the progress of the project.

Ulrike Baral
Alpnet Technology GmbH
TransAccount Project Co-ordinator