CENTRAL AND SOUTH-EAST EUROPEAN RESOURCES

CESAR

Central and South-East European Resources
Project no. 271022

Deliverable D1.2b
Periodic report after 18 months

Version No. 1.4
31/07/2012

Document Information

Deliverable number:	1.2
Deliverable title:	Periodic report after 18 months
Due date of deliverable:	31/07/2012
Actual submission date of deliverable:	31/07/2012
Main Author(s):	Tibor Pinter (HASRIL)
Participants:	Svetla Koeva (IBL) Maciej Ogrodniczuk (IPIPAN) Gyorgy Szaszak (BME) Marko Tadic (FFZG)
Internal reviewer:	Tamas Varadi (HASRIL)
Workpackage:	1
Workpackage title:	Management
Workpackage leader:	HASRIL
Dissemination Level:	Public
Version:	1.4
Keywords:	periodic report, 18 months

History of Versions

Version	Date	Status	Name of the Author (Partner)	Contributions	Description/ Approval Level
1.4		final	Tamás Váradi (HASRIL)	Supervision	
1.3		draft	Tibor Pinter (HASRIL)	corrections, editing	
1.2		draft	Svetla Koeva (IBL) Maciej Ogrodniczuk (IPIPAN) Gyorgy Szaszak (BME) Marko Tadic (FFZG)	contributions to WPs contributions to pms	
1.1		draft	Tibor Pinter (HASRIL)	core material	

Executive summary
The deliverable describes the work done from the 12th to the 18th months of the project. It contains an overview of the work completed and launched in the reporting period as well as major results and events. The deliverable also contains an update of pm-per-tasks indicator detailed by partners and WP-s.

Table of Contents

Table of Contents

1. Results of the third period described by WPs (13-18 months)

1.1 WP1 Management

- Reporting of the deliverables went on time. At the end of this period six deliverables are/were submitted.
- Budgetary issues after the first year.
- The web site (http://cesar.nytud.hu) was continuously developed and maintained.
- The internal project mailing list (cesar@nytud.hu) was maintained by HASRIL.
- Partners are involved in organizational tasks of the CESAR Road Shows.
- Tight cooperation with META-NET was maintained:
  - The coordinator kept regular, often daily contact with network manager Georg Rehm and held regular discussions with the coordinators of METANET4U and METANORD.
  - An active communication was held with META-NET Communication Group regarding presentation of META-NET at different events, cmpiling LWPs and defining META-NET visual appearance,.WP leaders were in tight contact with the respective people with other PSP projects
  - Active participation of other META-NET partners on CESAR Road Shows.
- Between M12 and M18 the following project meetings were held:
  - 16/04/2012 Brussels
  - 17/04/2006 Brussels
  - 20/06/2012 Brussels

Deliverables submitted in M13-M18:
- D1.2b Periodic report after 18 months (http://cesar.nytud.hu/deliverables/d1.2-periodic-project-reports-as-specified-in/d1.2b_periodic_report_18m.pdf)

1.2 WP2 Analysis and selection of language resources

- Partners have charted the national scene of their language community landscape.
- Partners have mapped the language service industry and language technology industry in their country.
- Partners have mapped the local policy makers in their country.
- Partners have identified and described language resources which are already developed or currently under development.
- The second release of the Contact database of stakeholders was made available.
- A report on available and potentially available resources was prepared.
- A report on resources of further interest was developed.
- CESAR partners gathered all relevant information (metadata) of the resources actually (or potentially) available to the consortium.
- From the early beginning of the project CESAR started to cooperate with META-NET (the cooperation related to works of WP2 meant synchronization with the templates, guidelines, specifications and available models for the resources and tools description provided by META-NET).
- CESAR supports the common and shared resource description between the four projects constituting METANET (i.e. CESAR, METANET4U, META-NORD and T4ME).
- Partners still worked on editing of the Language White Paper series.

Deliverables submitted in M13-M18:
- D 2.2b Contact database of stakeholders (http://cesar.nytud.hu/deliverables/d2.2-contact-database-of-stakeholders/1d2.2b_contact_database_of_stakeholders_18m.pdf)
- D 2.3c Report on resources (actually or potentially) available to the consortium (http://cesar.nytud.hu/deliverables/d2.3-report-on-resources-actually-or-potentially/d2.3c_report_on_resources_actually_or_potentially_available_to_the_consortium_m18.pdf)
- D 2.5: Report on resources of further interest (http://cesar.nytud.hu/deliverables/d2.5-report-on-resources-of-further-interest/d2.5report_on_resources_of_further_interest.pdf)

1.3 WP3 Enhancing language resources

Actions carried out in M13-M18:

- Partners presented their WP3-related partner activities at the project meeting in Brussels (April 16, 2012) and selected resourced at the demo session (April 17, 2012).
- META-NET representatives were contacted to review the current version of the schema and META-SHARE server at the META-FORUM meeting in Brussels (June 19-21, 2012).
- CESAR META-SHARE node located in Warsaw was upgraded to the newest version (2.1.2).
- New (batch 2) resources were gathered and processed by CESAR partners (concerning resource standardization, upgrade, alignment, documentation, IPR clarification etc.).
- Metadata descriptions of the new resources have been prepared in compliance with the newest META-SHARE metadata model.
- All previously uploaded metadata resource descriptions from batch 1 were converted to the newest version of the XML schema (2.1), reviewed and updated by partners and reuploaded to CESAR META-SHARE node.
- XSLT documentation generator has been upgraded to the newest version of the metadata model and used for generating XML-based human-readable resource descriptions. The current version of the generator was sent to META-NORD.
- New resources have been identified to be considered for the last batch.

Deliverables submitted in M13-M18:

- D3.2 – Second batch of resources (http://cesar.nytud.hu/deliverables/d3.2-second-batch-of-resources/d3.2_second_batch_of_resources.pdf)
- D3.2 Supplement to the second batch of resources (http://cesar.nytud.hu/deliverables/d3.2-second-batch-of-resources/d3.2.b_actions_on_resources.pdf)

1.4 WP4 Cross-national collaboration and Pilot service

- Installation of V2.1. META-SHARE software (node+metadata editor) at the official CESAR node in Warsaw was done
- Adaptation of V2.1. metadata schemes as well as conversion of schemas from batch 1 was made.
- Sharing of experiences with updated software and metadata (V2.1.) was done by partners.
- Continuous consultation on technical and legal questions and problems with META-NET and partner projects was run.
- Preparation for batch 2 was finished.
- Metadata descriptions and resources to the official CESAR node were uploaded.

Deliverables submitted in M13-M18:

- D4.4 Second upload of language resources (http://cesar.nytud.hu/deliverables/d4.4-second-upload-of-language-resources/d4.4_second_upload_of_language_resources.pdf)

1.5 WP5 Outreach, awareness and sustainability

- Preparations for organisation of CESAR road shows (Bulgarian, Slovak, Polish, Serbian, Croatian, Hungarian).
- Succesful organisation of CESAR road shows:
  - Bulgarian, Sofia, 2012-05-02
  - Slovak, Bratislava, 2012-06-07/08
- Successful dissemination efforts in EU and national level:
  - articles:
    - Ogrodniczuk, Maciej; Garabík, Radovan; Koeva, Svetla; Krstev, Cvetana; Pęzik, Piotr; Pintér, Tibor; Przepiórkowski, Adam; Szaszák, György; Tadić, Marko; Váradi, Tamás; Vitas, Duško. Central and South-European Language Resources in META-SHARE. // Infotheca. XIII (2012), 1; 3-26
    - Šojat, Krešimir; Mikelić Preradović, Nives; Tadić, Marko. Generation of Verbal Stems in Derivationally Rich Language. // LREC2012 Proceedings, pp. 928-933.
    - Berović, Daša; Agić, Željko; Tadić, Marko. Croatian Dependency Treebank: Recent Development and Initial Experiments. // LREC2012, pp.
    - "Towards a comprehensive open repository of Polish language resources" by Maciej Ogrodniczuk, Pęzik and Adam Przepiórkowski in Proceedings of the 8th International Conference on Language Resources and Evaluation (LREC 2012), pp. 3593–3597.
    - "PoliMorf: a (not so) new open morphological dictionary for Polish" by Marcin Woliński, Marcin Miłkowski, Maciej Ogrodniczuk and Adam Przepiórkowski in Proceedings of LREC 2012, pp. 860–864.
    - "The Polish Sejm Corpus" by Maciej Ogrodniczuk in Proceedings of LREC 2012, pp. 2219–2223.
    - "Central and South-European language resources in META-SHARE" by Maciej Ogrodniczuk, Radovan Garabík, Svetla Koeva, Cvetana Krstev, Piotr Pęzik, Tibor Pintér, Adam Przepiórkowski, György Szaszák, Marko Tadić, Tamás Váradi, Vitas Duško in INFOtheca: Journal of Informatics and Librarianship 13(1), 2012, pp. 3–26, ISSN 1450-9687.
  - conferences
    - LREC2012, Istanbul, 2012-05-21/27
      - two papers presentations as posters at the main conference
      - organisation of LREC2012 NooJ Tutorial
      - organisation of META-RESEARCH workshop on advanced treebanking (@META)
      - organisation of CESAR project appearance at the META-NET booth in EU Project Village.

- EAMT2012, Trento, 2012-05-28/30
  - poster and flyer project presentation
- NooJ2012, Paris, 2012-06-14/16
  - three papers presentations (two posters, one oral)
- META-FORUM2012, Brussels, 2012-06-20/21
  - poster and flyer project presentation
- project presentations
  - ICT-PSP Days, Zagreb, 2012-02-09
- Preparation of dissemination materials in coordination with META-NET such as flyer, posters, t-shirts, videos, cups, web-based material etc.
- Preparation of CESAR appearance at upcoming events and conferences (TSD, COLING,…)

Deliverables submitted in M13-M18:

- D5.4b Dissemination material (posters, flyers, brochures, t-shirts…) (http://cesar.nytud.hu/deliverables/d5.4-dissemination-material-posters-flyers/d5.4b_dissemination_material_m18.pdf)

1.6 Running tasks

- All tasks allocated for M18 are finished.

1.7 Delays in deliverables

- Deliverables were submitted in due time.

2. Task-person-months readiness overview

The overall work-flow can be measured on the person-months spent in the project. Below is a cumulative table of PM-s (with reference to the PM-state after M12).

	HASRIL	TMIT	FFZG	IPIPAN	ULodz	UBG	IPUP	IBL	LSIL	Total
Project effort - 12 months										
WP 1	12	1,25	3,4	0,8	1	1	2	1,5	1,3	**24,25**
WP 2	4,5	2,55	5,6	4,1	2,7	4	1	6	8,9	**39,35**
WP 3	19	61,2	12	41,40	16,5	14,5	30,1	21	14,3	**229,995**
WP 4	4	11,4	2,38	3,27	3	3,3	0,5	3,6	5	**36,445**
WP 5	8,5	3,5	11,5	1	1,5	6,5	1,5	3,2	4	**41,2**
Total	**48**	**79,9**	**34,88**	**50,56**	**24,7**	**29,3**	**35,1**	**35,3**	**33,5**	**371,24**
%	**41**	**93**	**39**	**43**	**44**	**40**	**45**	**35**	**60,7**	**48%**

	HASRIL	TMIT	FFZG	IPIPAN	ULodz	UBG	IPUP	IBL	LSIL	Total
Project effort - 18th months										
WP 1	20	1,64	5,4	1,2	2	1,5	3,2	2,5	2,4	**39,84**
WP 2	6	5	8,1	4,8	4	5	1,6	10	9,6	**54,1**
WP 3	31	99,9	22	91,69	29	29,5	44,4	34,5	20,8	**402,79**
WP 4	5	14,4	4,38	6,76	5	5	2,9	5	6,5	**54,94**
WP 5	15,5	8	21,98	3,81	5,5	7,3	7,6	8,7	7,7	**86,09**
Total	**77,5**	**128,94**	**61,86**	**108,26**	**45,5**	**48,3**	**59,7**	**60,7**	**47**	**637,76**
%	**66**	**151**	**70**	**65**	**81**	**66**	**77**	**61**	**80**	**77%**

2.1 Submitted deliverables after 18 month

D1.1 Internal and external project website (3)

D1.2 Periodic project reports as specified in Appendix X (6)

Table of person months progress on deliverable readiness after 6 months

D1.2b Periodic project reports as specified in Appendix X - after M18 (18)

D2.1 Language Whitepaper:

(English version): Bulgarian, Croatian, Hungarian, Polish, Slovak, Serbian (6)

(localized version): Bulgarian, Croatian, Hungarian, Polish, Slovak, Serbian (6)

Compressed version

D2.2 Contact database of stakeholders (pdf, xls) (4)

D2.2b Contact database of stakeholders - after M18 (18)

D2.3 Report on resources (actually or potentially) available to the consortium (6)

D2.3b Report on resources (actually or potentially) available to the consortium (12)

D2.3c Report on resources (actually or potentially) available to the consortium (18)

D2.4 Report on methodology and criteria for the selection of resources (8)

D2.5 Report on resources of further interest (18)

D3.1 First batch of resources complying with the project's technical, linguistic, legal requirements (10)

Supplement for the D3.1 (10)

D3.2 Second batch of resources (18)

Supplement for the D3.2 (18)

D3.3 Third batch of language resources (24)

D4.1 Metadata descriptions and other interoperability standards as agreed with META-NET and partner projects (3)

D4.2 IPR principles and provisions as agreed with META-NET and partner projects (4)

D4.3 First upload of language resources (10)

D4.4 Second upload of language resources (18)

D4.5 Third upload of language resources (24)

D5.1 Action plan, Updated Action plan (4, 12)

D5.2 Awareness, mobilization and dissemination actions - described in the annual reports (12, 24)

D5.3 Sustainability strategy and plans beyond the end of the project (24)

D5.4 Dissemination material (posters, flyers, brochures, t-shirts…) (6)

D5.4b Dissemination material (posters, flyers, brochures, t-shirts…) - after M18 (18)

3. Consortium meetings

In the last 6 months the consortium held several meetings on an occasional basis. As the ongoing work went without any special problem, face-to-face meetings were limited to the minimum:

1. Brussels - 16/04/2012
2. Brussels - 17/04/2012
3. Brussels - 20/06/2012