

PROJECT PERIODIC REPORT				
Czech-Norwegian Research Programme (CZ09)				
Norwegian Financial Mechanism 2009-2014				
Programme area	Bilateral Research Cooperation			
Periodic report	1/2014			
Period covered	1 October to 31 December 2014			
Project ID number	7F14047			
Acronym	HaBiT			
Project title in English	Harvesting big text data for under-resourced languages			
Project title in Czech	Získávání velkých textových dat pro jazyky s nedostatečným množstvím jazykových zdrojů			
Project Promoter (name, full address)	Masarykova univerzita Žerotínovo nám. 617/9, 601 77 Brno Czech Republic			
Project Partner(s) (name, full address)	Norges teknisk-naturvitenskapelige universitet Høgskoleringen 1, 7491 Trondheim Norway			
Full Name of Principal Investigator	doc. PhDr. Karel Pala, CSc.			
Signature of Principal Investigator				
Statement	<i>I hereby declare that the information I state in the project periodic report is accurate, true and complete. I am aware that if the information has been reversed in the opposite, I will face sanctions from the Programme Operator.</i>			
Done in	Brno, Czech Republic			
Date	10/01/2015			
On behalf of Project Promoter				
Stamp of Project Promoter				
Statutory authority of Project Promoter	Name(s):			
	Signature(s):			
	Position:			

## 1. GENERAL INFORMATION ABOUT PROJECT

### 1.1 Activity of research and development in project

Basic research <input checked="" type="checkbox"/> 0-100 %	Applied research <input type="checkbox"/> 0-100 %	Experimental development <input type="checkbox"/> 0-100 %
---	--	--

#### Project Promoter: Masaryk University

Basic research <input checked="" type="checkbox"/> 0-100 %	Applied research <input type="checkbox"/> 0-100 %	Experimental development <input type="checkbox"/> 0-100 %
---	--	--

#### Project partner: Norwegian University of Science and Technology

Basic research <input checked="" type="checkbox"/> 0-100 %	Applied research <input type="checkbox"/> 0-100 %	Experimental development <input type="checkbox"/> 0-100 %
---	--	--

### 1.2 Official project starting date reported to Programme Operator

01/10/2014
------------

### 1.3 Project duration in months in total

31
----

### 1.4 Total project cost in Project contract (in CZK)

24,468,000
------------

#### 1.4.1 Total grant in Project contract (in CZK)

24,468,000
------------

#### 1.4.1.1 Grant for reporting period approved by Programme Operator (in CZK)

1,892,000
-----------

### 1.5 Number of partners (including Project Promoter)

2
---

### 1.6 Ethical issues

NO
----

## 2. SCIENTIFIC AND MANAGEMENT PART

### 2.1 Publishable summary in English

The main objectives of the HaBiT project are to gather large-scale text data (corpora) from the Web for under-resourced languages, involving Norwegian, partly Czech and the major languages of Ethiopia — Amharic, Afaan Oromo, Tigrinya, Somali — and to build shallow processing applications. The gathered data will be processed to make it usable in many language applications, such as information extraction or machine translation. Furthermore, in the process of collecting corpora data, existing tools for building web text resources will be further developed and improved since the Ethiopian languages are quite different from most European languages. Applications for the given languages will be built to allow for the separation and disambiguation of multiple senses of words.

One of the project's aims is to build a multi-billion word Norwegian corpus using the tools co-developed by the FI MU NLP Centre, Masaryk University and utilized in a previous joint EU-funded project with NTNU ("PRESEMT: Pattern REcognition-based Statistically Enhanced MT", 2010-2012). Second, NTNU collaborates with University of Oslo and two Ethiopian universities in a project to support linguistic resource building in Ethiopia funded by Norad ("Linguistic Capacity Building — tools for the inclusive development of Ethiopia", NORHED 2013-2018). It is natural to link these activities and to include processing of the four major languages in Ethiopia in the present project: the HaBiT project will be able to feed into and leverage on the NORHED project, thoroughly testing the technologies and thus addressing also the call topics on technology assessment, verification and testing, as well as on ICT meeting societal challenges, hence obtaining a relevant added value also in the political respect through cooperation with a less-developed country.

In the first months of the project, the main attention has been paid to the specification of the corpora and corpus building modules, and the requirements for word-sketch grammars and tools, that will be used for gathering data.

Up-to-date information can be found at the project website [www.habit-project.eu](http://www.habit-project.eu).

#### 2.1.1 Publishable summary in Czech

Hlavním cílem projektu Habit je jednak shromáždit velká textová data (korpora) z Webu pro jazyky s nedostatečnými zdroji včetně norštiny a částečně češtiny a také hlavní etiopské jazyky - amharštinu, afaan oromštinu, tingrinštinu, somálštinu a jednak vytvořit aplikace pro jejich mělké zpracování. Získaná data budou zpracována tak, aby mohla být použita v řadě jazykových aplikací, jako je extrakce informací a strojový překlad. V průběhu budování korpusových dat budou dále vyvíjeny a zlepšovány existující nástroje pro získávání webových textových zdrojů, protože etiopské jazyky se výrazně liší od většiny evropských jazyků. Aplikace pro zmíněné jazyky budou budovány tak, aby dovolily zkoumat a vyčleňovat víceznačnost slovních významů.

Jedním z cílů projektu bude vytvořit více miliardový korpus norštiny, a to s použitím nástrojů spoluvyvinutých Centrem ZPJ FI Masarykovy univerzity a využitých v předchozím společném EU projektu realizovaném spolu s NTNU ("PRESEMT: Pattern REcognition-based Statistically Enhanced MT", 2010-2012). Za druhé, NTNU spolupracuje s Univerzitou v Oslo a dvěma etiopskými univerzitami v projektu podporujícím budování jazykových zdrojů v Etiopii a finančně zajištěným organizací Norad ("Linguistic Capacity Building – tools for the inclusive development of Ethiopia", NORHED 2013-2018).

Je přirozené propojit tyto aktivity a zahrnout zpracování čtyř hlavních etiopských jazyků do současného projektu: HaBiT project tím umožní datové vazby na NORHED projekt a důkladné testování technologií. Tak bude možno adresovat rovněž témata týkající se hodnocení technologií, jejich verifikování a testování včetně toho, jak informační technologie (ICT) splňují sociální výzvy. To umožní získat relevantní přidanou hodnotu také v politickém ohledu díky spolupráci s méně rozvinutou zemí.

V prvních měsících projektu se hlavní pozornost zaměřuje na specifikaci textových korpusů a na modul pro jejich budování, dále pak na požadavky ohledně gramatik pro slovní profily a nástroje, jichž se užívá pro sběr textových dat. Aktuální informace lze najít na webové stránce projektu: [www.habit-project.eu](http://www.habit-project.eu).

## 2.2 Project objectives for reporting period

The main efforts during the first months of the project have been focused on the technical part of the data gathering. Detailed specifications of the corpora have been prepared in order to collect the right resources from the Web. Corpus building modules, including the tool “unitok”, have been designed and developed by the MU team for this purpose. The raw web data contains boilerplate data, that is not suitable for research purposes, but the developed tools enable the processing of the data into formats more appropriate for a representative text corpus. The tool “unitok” is publicly available under the GPL licence at <http://corpus.tools>.

Using these corpus building modules, the first version of a corpus for the Amharic language was collected from the Web, having about 10 million tokens, which is approximately 9 million words. This dataset has been made available to the NORHED colleagues in Ethiopia for feedback and will be also used as a base for testing the implemented tools.

During the last weeks of year 2014, the first requirements for word-sketch grammars were defined. These will be completed by the end of January 2015.

## 2.3 Work progress and achievements during reporting period

For this reporting period, all objectives given in the project plan have been successfully achieved, taking into account the delayed project start.

A first version of the Amharic corpus was successfully built. The cleaned and de-duplicated version contains almost 9 million words, which is about a half of the targeted final version. We consider these results preliminary so far, the full version is planned to be completed in the ninth month of the project. Access to the Amharic corpus has already been given to the NORHED Ethiopian colleagues; feedback is awaited and all possible changes will be incorporated into the data.

Work in WP1 has just started and the planned end has not been achieved yet, but two deliverables from this WP are in progress:

- D1.1.2 Specification of the corpora and the corpus building module
- D1.1.3 Specification of word-sketch grammars and tools

These documents will be evaluated and completed in the sixth month of the project.

### 2.3.1 Description of project implementation

#### Work Package 1

A first version of the specification of the corpora was drafted during the first three months of the project. A detailed analysis of the resources available on the Web is being prepared and will be used when crawling the Web to obtain new corpora for the under-resourced languages.

Preparation of the corpus building module has been started and will be finished within the first six months of the project.

#### Work Package 3

WP3 was planned to start only after M6, but it was possible to start it already in M3: Crawling text data for Amharic resources started after the analysis in WP1. The tool SpiderLing was used for data mining and the resulting raw data from the Web was then further refined and processed. First by the tool “jusText”, to remove boilerplates, and then by the text tokenization tool “unitok”. The first version of the corpus now has approximately 10 million tokens and has been indexed for fast searching.

As described above, this corpus version has already been made available to our Ethiopian colleagues, who will send us feedback, which will be incorporated when preparing the second version of the corpus.

### 2.3.2 Milestones achievement

No.	Milestone name	WP no.	Lead partner (abbreviation)	Planned achievement date dd/mm/yyyy	Actual/Forecast achievement date dd/mm/yyyy
1	MS1 Kick-Off Meeting		MU	31/10/2014	21/11/2014
2	MS2 System specification	1	MU	31/03/2015	28/02/2015

### 2.3.3 Deliverables achievement

No.	Deliverable name	WP no.	Lead partner (abbrev.)	Planned delivery date dd/mm/yyyy	Actual/Forecast delivery date dd/mm/yyyy
1	D1.1.2 Specification of corpora and the corpus building module	1	MU	31/03/2015	28/02/2015
2	D1.1.3 Specification of word-sketch grammars and tools	1	MU	31/03/2015	28/02/2015

## 2.4 Work Packages (WPs)

### 2.4.1 Project working packages (WP)

WP number	Title	Planned date of start (dd/mm/yyyy)	Actual/Forecast date of end (dd/mm/yyyy)
WP1	System integration	10/2014	04/2017
WP3	Corpora for under-resourced languages	03/2015	11/2016

#### 2.4.1.1 WP number

WP1

#### 2.4.1.2 WP title

System integration

#### 2.4.1.3 WP leader

Karel Pala

#### 2.4.1.4 WP start date

01/10/2014

#### 2.4.1.5 WP end date

30/04/2017

#### 2.4.1.6 WP objective

- To define the overall system specifications
- To integrate and test all software modules produced in WP4-WP6 in an iterative way
- To create a demo website to illustrate the project to end users and potential stakeholders
- To create the overall integrated system

#### 2.4.1.7 WP task

##### PLAN:

##### Task T1.1: System and modules requirements specification

- Overall system specifications; definition of the functional and non-functional requirements.
- Specifications of the modules to be developed in WP4, WP5 and WP6.

##### Task T1.2: Module integration and testing

- Integration of the modules developed in WP4, WP5 and WP6 into the HaBiT system prototype.
- Prototype testing and evaluation in three iterations for software review and upgrade.

##### Task T1.3: Coordination of the platform development

- To guarantee that the modules are developed in a consistent way and in line with the system specifications.
- Cross-WP validation of the software components.
- Create the final system demonstrator.

#### 2.4.1.8 WP deliverable

##### PLAN:

**D1.1.1** System specifications [M6]: Overall system design definitions

**D1.1.2** Specification of corpora and the corpus building module [M6]

**D1.1.3** Specification of word-sketch grammars and tools [M6]

**D1.1.4** Specification of the semantic content matching and wordspace module [M6]

**D1.2.1** The HaBiT system v1 [M14]: First integrated system prototype

**D1.2.2** The HaBiT system v2 [M20]: Second integrated system prototype

**D1.2.3** The HaBiT system v3 [M29]: Third and pre-final integrated system prototype

##### ACHIEVEMENTS:

Two deliverables are being prepared: D1.1.2 'Specification of the corpora and the corpus building module' and D1.1.3 'Specification of word-sketch grammars and tools' and are expected to be finalized in Feb 2015.

#### 2.4.1.9 WP milestone

The MS1 milestone has been achieved in the form of a kick-off meeting. The meeting was held on Nov 20-21 2014 in Brno with the participation of both partners.

#### 2.4.1.10 WP Human resources

Qualification level: 9 - Karel Pala (docent), Aleš Horák (docent), Pavel Rychlý (PhD), Miloš Jakubiček (RNDr., PhD student), Vojtěch Kovář (RNDr., PhD student), Björn Gambäck (professor), Janne Bondi Johannessen (professor).

Person-months: 3

#### 2.4.2 Project output(s)

Type of output	Title	Date of accomplishment (mm/yyyy)	Date of realization (mm/yyyy)
D	Finding the Best Name for a Set of Words Automatically	11/2014	12/2014
D	Intelligent Search and Replace for Czech Phrases	11/2014	12/2014
D	Mapping Czech and English Valency Lexicons: Preliminary Report	11/2014	12/2014
D	Text Tokenisation Using unitok	11/2014	12/2014



## 2.5 Project management during reporting period

The project consortium held a kick-off meeting on November 20th and 21st, 2014 in Brno, Czech Republic. The Project Management Board was established; its members are Karel Pala from MU and Björn Gambäck from NTNU. These members can be substituted by WP leaders when appropriate. The main points discussed during the kick-off meeting were important changes concerning the budget and the shortened project period. Also, a short overview of the project contract ("Rozhodnutí") was prepared for the Norwegian partners as it is not yet available in English. Responsibilities and consequences were explained to partners to keep them in the picture, however, the fact that the English translation of the Rozhodnutí has not still been prepared by the MŠMT makes the cooperation with Norwegian partner complicated.

The MU team proposed to use the Trac system for workflow communication in the project, it is a complex system for project management and supervision of milestones with a built-in ticket system, so all team members are aware of what is being in progress and the WP leaders can easily monitor if the given goals can be reached on time.

Internal communication takes place through the project website [habit-project.eu](http://habit-project.eu) via access for authorised team members. The public part of the website offers basic information and is being developed according to the Communication Plan and project manual.

The consortium members agreed on periodical video-conference meetings, so any major issues can be discussed on time and the necessary plans can be created. Any other communication will be made through e-mail aliases, which will ensure that the whole team — either MU or NTNU — will be aware of the project progress.

There are no changes in the consortium, nor in the legal statutes of the participating organisations. However, the project plan has had to be revised due to the delays in the funding decision and in the contract signing process, as well as in the communication of the envisioned revised project start date and project period. In effect, almost all milestones and deliverables have been moved forward in time and the project plan been compress in order to account for the shortened project time. The fact that the envisioned project start was only conclusively communicated to the project partners at the end of November 2014 further delayed the project work, even though some efforts still could be carried out from October 1. In particular, the delay made it impossible to retain the planned staffing at NTNU, as further described below (see Section 2.3.5).

### 2.5.1 List of Project Promoter's staff working on project during reporting period (all members connected to personal costs) and changes at team

Full name	Position in project	Full-/part-time	Work load (1.0-0)	Hired on (dd/mm/yyyy)	Quit on (dd/mm/yyyy)
doc.PhDr. Karel Pala, CSc.	project team leader	part-time	0.20	01/10/2014	--
doc.RNDr. Aleš Horák, Ph.D.	researcher	part-time	0.20	01/10/2014	--
Mgr. Pavel Rychlý, Ph.D.	researcher	part-time	0.20	01/10/2014	--
Mgr. Vít Suchomel	researcher, PhD student	part-time	0.40	01/10/2014	--



Mgr. Vít Baisa	researcher, PhD student	part-time	0.25	01/10/2014	--
Mgr. Miloš Jakubíček	researcher, PhD student	part-time	0.25	01/10/2014	--
RNDr. Vojtěch Kovář, Ph.D	researcher	part-time	0.25	01/10/2014	--
Bc. Lucia Kocincová	project administrator	part-time	0.30	01/10/2014	--
<b>Staff changes</b>					
--	--	--	--	--	--

*Full-time work means working only on the project.*

## 2.5.2 List of Project partner's staff working on project during reporting period (all members connected to personal costs) and changes at team

Full name	Position in project	Full-/part-time	Work load (1.0-0)	Hired on (dd/mm/yyyy)	Quit on (dd/mm/yyyy)
Prof. Björn Gambäck, PhD	partner team leader	part-time	0.20	01/11/2014	
Prof. Janne Bondi Johannessen, PhD	researcher	part-time	0.10	01/11/2014	
Partha Pakray, PhD	researcher post-doc	full-time	1.0	01/03/2015 *	
NN, PhD	researcher post-doc	full-time	1.0	01/03/2015 *	
<b>Staff changes</b>					

\* Envisioned start dates.

## 2.5.3 Personnel changes justification

The original staff plan at NTNU included Dr. Partha Pakray in a 2-year postdoctoral position from Month 3 of the project and a PhD student from the start of the project. However, the shortened project time has made it impossible to employ a PhD student exclusively on this project. Furthermore, the delay in the project start severely affected Dr. Pakray's involvement in the project: at the time of the submission of the project application, he was employed at NTNU, but he had to leave Norway when that employment contract ended on June 30, 2014, without the HaBiT project having been confirmed at the time. Since the project was only confirmed towards the end of November 2014, the process to properly announce the vacancy had to be postponed until then — and the process to obtain a new working permit visa for Dr. Pakray could only be initiated at the beginning of 2015.

The effect of these two changes will be that the Norwegian team rather will aim to employ two postdoctoral researchers to work full-time on the project during two years each. One of the two postdocs positions will be filled by Dr. Pakray who is expected to be able to join the project from the beginning of March 2015 and thus work until the end of February 2017, while the other postdoc position will be filled by a candidate joining the project in May 2015 and working until the project's end date (April 30, 2017).

## 2.6 Monitoring and auditing of project

Monitoring of the technical content and progress is run by the Project Management Board, which consists of the team leaders of each party. The Board is responsible for handling most of the scientific issues, planning of the project, and changes of the work plan. In case of any important issues or problems, the Project Management Board has the responsibility to resolve them. Moreover, the Board reviews all publications and project deliverables.

Administrative issues, including day-to-day operations, and regular project management are handled by the Project Promoter.

The project has not been audited from any institutions.

Each of project partners proposes an independent expert, who will evaluate the project report and its procedures. On the MU side the expert is Prof. Václav Matoušek from the West Bohemia University in Pilsen. On the NTNU side the proposed expert is Prof. Lars Borin from the University of Gothenburg.

### 2.6.1 Risk management and quality assurance

Quality of each project component and risk management will be monitored and evaluated according to a Quality Assurance Plan. This document is in draft form at the moment of reporting, since there are several issues that need to be clarified by the Ministry and by both project partners in order to consider the Plan finalized.

Risk assessment is controlled by the project promoter, who ensures that this process will be continuous throughout the whole project duration. The project promoter also discusses relevant risk issues with the Board at regular meetings to secure that each potential risk is taken seriously. Such issues will be also part of Risk Management Plan which is being drafted and will be finished within first six months of the project, when the first important milestone should be reached. The Risk Management Plan will assess any potential probabilities of risks and will also define measures that will be needed to be taken in order to minimize any impact of the risks. This plan will be reviewed and improved regularly to ensure that any new risks are taken into account.

### 2.6.2 Irregularities

No irregularities have occurred.

## 2.7 Intellectual property rights management

The main aim of the consortium is to make any project output — resources and tools — open access. Access rights will be given to all the project partners including the Ethiopian collaborators. The Project Management Board controls the distribution of intellectual property rights in accordance with the Partnership Agreement. All decisions regarding intellectual property rights will be summed up in the Technology Management Plan, which will also contain plans for the exploitation of results and licensing strategies.

## 2.8 Scientific publications and dissemination of project in report period

The initial progress of the project has been published at the Eighth Workshop on Recent Advances in Slavonic Natural Language Processing (RASLAN 2014), which was held in Karlova Studánka in the Czech Republic. The MU team published four papers directly related to the HaBiT project; more information can be found below.

### 2.8.1 Reporting on scientific publications

- Finding the Best Name for a Set of Words Automatically*** - Pavel Rychlý

Many natural language processing applications use clustering or other statistical methods to create sets of words. Such sets group together words with similar meaning and in many cases humans can find an appropriate term quickly. On the other hand computers represent such sets with a meaningless number or ID. This paper proposes an algorithm for automatic finding of names of word sets. It provides result examples as a simple evaluation of the method.
- Intelligent Search and Replace for Czech Phrases*** - Zuzana Nevěřilová, Vít Suchomel

This work proposes a new improvement of the 'Search and Replace' function well known from most text processing software. The standard search and replace function is used to replace exact form of words or phrases by another words or phrases in text documents. It is quite sufficient for languages with minimal inflection such as English. However, a well working word or phrase replacement function for morphologically rich languages requires much more thought. We explore the issues of implementing a useful search and replace in the Czech language and propose solutions to majority of the problems: A syntactic parser is employed to identify the phrases containing the search word or phrase. The correct word forms used as a replacement are generated by a morphological analyser. A web demonstration utilizing the proposed solution is presented. The attached examples of use reveal the cases in which the implemented method works well.
- Mapping Czech and English Valency Lexicons: Preliminary Report*** - Karel Pala, Vít Baisa, Zdeňka Sitová, Jakub Vonšovský

We describe here a very first attempt to connect two valency lexicons: Pattern Dictionary of English Verbs (PDEV) and VerbaLex. Both lexicons contain verbs together with their syntactic structure (arguments of the verbal predicate) and semantic restrictions (semantic types typical for a given verb argument). The lexicons are similar in overall but differ in details since their formalisms are tailored for the respective languages. They also differ in a way they have been built: whilst the former resource has been built using Corpus Pattern Analysis methodology the latter has been built upon previous datasets Brief, Vallex and Czech WordNet. We present a preliminary work on linking English patterns in PDEV with their Czech equivalents: frames in VerbaLex.
- Text Tokenisation Using Unitok*** - Vít Suchomel, Jan Michelfeit, Jan Pomikálek

This paper presents unitok, a tool for tokenisation of text in many languages. Although a simple idea – exploiting spaces in the text to separate tokens – works well most of the time, the rest of observed cases is quite complicated, language dependent and requires a special treatment. The paper covers the overall design of unitok as well as the way the tool deals with some language or web data specific tokenisation cases. The rule what to consider a token is briefly described. The tool is compared to two other tokenisers in terms of output token count and tokenising speed. unitok is publicly available under the GPL licence at <http://corpus.tools>.

## 2.9 Project promotion and information activities about project

During the first three months of the project, an analysis of possible activities and outputs was prepared in both countries. This analysis contains many options that will be evaluated and incorporated in the publicity plan.

The main focus and communication platform will be project website, which is under development. Within the first six months of the project an up-to-date website with information for each group of possible audiences will be prepared.

Both partners organize promotion actions for the general public, so such potential events will be taken into account and presentations of the project will be prepared. An informative text for the Masaryk University journal MUNI is going to be prepared in the first half of 2015. A joint meeting with the Ethiopian partners is planned for September 2015 (in connection with a meeting in the NORHED-financed project in Oslo), in order to further liaison with them and to promote the HaBiT project results.

### 2.9.1 Project website

[www.habit-project.eu](http://www.habit-project.eu)

## 2.10 Achievement of Programme outcome(s) and outputs

The HaBiT project contributes to research-based knowledge development by extending the cooperation within and outside of the Czech-Norwegian Research Programme. A number of PhD students are already part of the research and the project partners plan to include more female researchers, PhD students, and postdocs in order to actively engage them in best practices of both workplaces, and to transfer and share state-of-the-art development.

This close cooperation between the partners can become a basis for future successful cooperation, thus ensuring sustainability of the research or other future relevant activities.

The project results will also contribute to the cultural development of less-developed countries by acquiring and sharing developed information technologies.

All project outputs will be presented to the international research community and will also be disseminated via the project website in order to present the research to all possible types of audiences.

### 2.10.1 Programme output indicators in project (numbers)

Number of PhD students	3
Number of postdocs	2
Number of female researchers (including after maternity leave)	2
Number of internationally refereed (joint) scientific publications	0

*Note: Numbers of the target groups are for the whole project. These indicators are used for reporting about the Programme.*

### 3. FINANCIAL PART

#### 3.1 Explanation of use of resources

The expenditures at Masaryk University have been spent in agreement with the proposal in the following structure:

- personnel costs - 384 659,- CZK - part-time contracts of 8 researchers
- travel costs - 23 548,- CZK - RASLAN workshop active participation of 3 researchers (Pala, Rychlý, Suchomel)
- indirect costs (overheads) - 244 324,- CZK - computed as 60% Flat rate

Due to the delayed start of the project the expenditures have been considerably lower than it was expected (3 months instead of 7 months). As detailed in Section 2.5, this in particular affected the resource usage on the Norwegian side: since all personnel involved from Norway so far have been on permanent academic positions, no personnel costs directly related to the project have been claimed for the reporting period. The only costs incurred by the Norwegian partner relate to Gambäck's participation in the project kick-off (also Pakray participated physically in the kick-off meeting, but he was already in Brno at the time; Johannessen participated remotely in the kick-off meeting over a Skype connection).

#### 3.2 Budget changes justification for reporting period

The only budget changes are related to the delayed beginning of the project, i.e., the project expenditures correspond to a 3 months period instead of 7 months as planned.

#### 3.3 VAT reclaim YES or NO

Name of Project promoter/partner	VAT reclaim (Yes/No)
Masaryk University	NO
Norwegian University of Science and Technology	NO

#### 3.4 Indirect costs model - overheads

Name of Project Promoter/Project partner (abbreviation)	Participant Identification Code (PIC)	Overheads rate in % (per partner)	Using analytical accounting system (Yes/No)
Masaryk University	999880657	60	YES
Norwegian University of Science and Technology	999977851	60	YES

#### 3.5 Procurement and small scale contracting

No procurement or small scale contracting are planned for the project, so none were launched in the reporting period.

#### 3.6 Fund for bilateral activities

No bilateral funds received.

## 4. MANDATORY AND VOLUNTARY ANNEXES

### 4.1 Overview of annexes required to project periodic report

No.	Annex	Mandatory online submission format
I.	Actual incurred expenditure in reporting period (signed by statutory of the Project Promoter) Annex I (template)	xls(x)/pdf
II.	Proposed expenditure in following reporting period (signed by statutory of the Project Promoter) Annex II – on later request by PO	xls(x)/pdf
III.	Report on actual incurred expenditure by Czech institutions (Výkaz skutečně vynaložených výdajů dle zvyklostí instituce)	pdf
IV.	Form C (FP7) based form on financial reporting – Norwegian partners (template)	pdf
V.	Reports by periodic report evaluators (min. 2 oponentní posudky) (template)	pdf
VI.	Letter of attorney, if applicable (plná moc zplnomocněné osoby)	pdf
VII.	Photo documentation (voluntary)	CD/pdf

## 5. OTHER ANNEXES

N/A