MOSES CORE

Deliverable D2.2

The 2013 Machine Translation Marathon (MTM13)

Work Package:	WP2: Machine Translation Marathons
Author(s):	Ondřej Bojar, Achim Ruopp
Due Date:	January 31st, 2014
Dissemination Level:	Public

June 11, 2014

1 Aim

This report describes the 2013 Machine Translation Marathon (MTM13) as organized by the Institute of Formal and Applied Linguistics (ÚFAL), Charles University in Prague, and held in Prague in September 9–14, 2013.

The aim is to summarise the activity of the MTM, and report on the survey conducted after the Marathon using an online feedback form to learn for future Marathons.

The full text of the feedback form is available as Attachment A followed by a detailed summary of the responses in Attachment B. Abbreviations used occassionally in this report are listed in Attachment C.

2 Introduction

The 2013 marathon was the eighth in the series, and the second one sponsored by MosesCore.

The marathons traditionally mix introductory lectures and labs for newcomers, advanced research talks and, most importantly, projects. The overall aim is to foster the development and use of open source MT software.

The target audience of MT Marathons are MT developers, researchers and users. At MTM13, we specifically targeted also participants from industry, ranging from managers to software developers.

There are four main parts to the MT Marathon:

- collaborative hacking projects,
- the open source convention, i.e. presentation of papers on new open-source tools for MT,
- the "summer" school with lectures and labs given by leading researchers in the field,
- invited talks on current MT-related topics.

The marathon in Prague had three extra items:

- round table on open source MT for commercial use, organized by TAUS,
- round table on best practice with managing research computer clusters,
- a tutorial on running Moses on Amazon EC2.

The MTM13 website is available at `www.statmt.org/mtm13` including the programme with links to video recordings wherever available.

3 The MT Marathon

3.1 Participation

The first call for participation in MTM13 was issued in May and aside from public announcements, we also specifically sent out invitations to a few (mainly Czech) companies that are interested in or already use MT.

By the time the event started we had 113 registered participants (including invited speakers and lecturers and three on-site registrations but excluding a few extra participants who arrived without registering). Only 10 of the registered people did not make it to Prague in the end. In total, we had 103 attendees. We were pleased to see such a low rate of "no-shows", given that registration is free.

From the affiliations provided by the attendees, 63 (61%) were academics, with the rest coming from commercial organisations or government bodies. This result confirms that we were successful in attracting participants outside of academia: 39% this year compared to 23% in 2012.

The feedback form results (submitted by 45 attendees) provide a finer detail on participants: 36% came from industry (with slightly more management-level employees than researchers or developers), 24% were postgraduate students and 24% were researchers in academia, see Appendix B for the complete listing.

Since MosesCore provided funding to support students from institutions outside the consortium, we issued a call for bursary applications, with funding of €500 each available for up to 6 students. These were intended to support those relatively new to MT, and applications had to be supported by short statements from the student and supervisor. There were exactly 6 applications and since they all satisfied the criteria, they were all accepted. In the end, 5 of the students were able to attend and receive the bursaries. The funds allocated for the unclaimed bursary were left unspent.

3.2 Projects

MTM open source projects are week-long hacking sessions, conducted in small groups formed on the first day, and aiming to implement or extend open source MT software, or to try out a new research idea. For those more experienced in the field, projects are the main business of the MTM.

We followed last year's good experience and collected project proposals in advance. What proved particularly useful and at times almost interactive, was to use a shared online list of project proposals where anybody could have contributed any text. This document both in its editable version as well as

a snapshot in PDF are available from the corresponding MTM13 web page.[1]

The actual project groups were formed on Monday, after each proposer presented his or her project. What worked particularly well and was also positively mentioned in the feedback forms, was to use the blackboard where project leaders indicated where they are waiting for prospective team members and everybody marked with a simple tick their interest in the various projects. This allowed project leaders to see if their team is likely to get sufficiently big, and perhaps also contributed to some "load balancing" since everybody saw which projects are going to be crowded.

There were 19 projects announced on the first day of the Marathon and 14 made it up to the last day, delivering a brief summary on Saturday. The slides for all project sessions (boaster session on Monday, interim reports on Wednesday, final reports on Friday) are available in MTM13 SVN repository and linked from the programme web page[2]. Here are project titles from the final presentations:

- CorefMT (2 members)
- Forest MIRA (Forest rescoring in Joshua for MIRA training; 2 members)
- Inline Tag Handling (5 members)
- Internal tree structure for GHKM rules in Moses (4 members)
- Language Model Interpolation (2 members)
- Jacana Word Aligner (1 member)
- New features, testing and refactoring Joshua (2 members)
- A Discriminative Lexicon for Translating to Morphologically Rich Languages (10 members)
- MTSpell (5 members)
- Multipass Decoding in Moses with CSLM (4 members)
- Extending KenLM Pruning (3 members)
- QuEst@MTM (11 members)
- Sparse Features for Reordering (3 local and one remote member)

[1] `http://ufal.mff.cuni.cz/mtm13/projects.html`

[2] `http://ufal.mff.cuni.cz/mtm13/programme.html`

- Social Media Machine Translation Toolkit (SMMTT; 4 members)

As last year, we also had a project initiated and led by a person coming from industry: Inline Tag Handling led by Achim Ruopp.

Based on the feedback form, both experienced researchers and newcomers recognize that projects are the cornerstone of MT Marathon. A few people mentioned they were sorry for having decided to skip projects, because they wanted to focus on other activities. This is in line with several suggestions to cut down on other activities and distractions in favor of projects. One response mentioned that there were so many good talks this year that it, ironically, had a bad impact on time available for projects. A related remark promoting projects said that other activities, e.g. the various round tables, distract experienced people from contributing to their project teams.

The project presentations (midweek and final reports) were also well perceived; only 4% of the respondents deliberately skipped them. Specific comments in the responses highlight the importance of the slides as a minimalistic documentation and pointers to project participants for the future. If possible, MT Marathon organizers should remind project members to notify the community (e.g. in the Moses mailing list) in case a project evolves further, after the MT Marathon.

Since the full 100% of responses said that projects should be kept in, we indeed suggest to try cutting down on other activities a little, giving more space for project work without the need to work till late or very late. On a less serious note though, we feel that there is a reason why this event is called a "marathon".

3.3 Open Source Convention: Papers

The call for papers asked for submissions describing new open source MT software, and extensions to existing tools. This call gives MT researchers and developers the opportunity to share information about implementation, and to publicise their software – an opportunity which is generally not available at typical research conferences. The accepted papers are published in the Prague Bulletin for Mathematical Linguistics (PBML)[3].

We received 17 submissions for MTM13 and after two independent reviews, 16 were accepted for publication in PBML and presentation at the MTM. Of these, 10 were selected for publication in Volume 100, printed and made available at MTM13. The remaining 6 papers are scheduled for Volume 101 due April 2014. This division is mainly driven by physical constraints of

[3] `http://ufal.mff.cuni.cz/pbml`

the printed version of PBML but it allowed us to provide more space to the articles that deserve it, at the cost of a later publication date.

The accepted papers were:

- *CASMACAT: An Open Source Workbench for Advanced Computer Aided Translation* by Vicent Alabau, Ragnar Bonk, Christian Buck, Michael Carl, Francisco Casacuberta, Mercedes García-Martínez, Jesús González, Philipp Koehn, Luis Leiva, Bartolomé Mesa-Lao, Daniel Ortiz, Herve Saint-Amand, Germán Sanchis, Chara Tsoukala
- *COSTA MT Evaluation Tool: An Open Toolkit for Human Machine Translation Evaluation* by Konstantinos Chatzitheodorou, Stamatis Chatzistamatis
- *DIMwid – Decoder Inspection for Moses (using Widgets)* by Robin Kurtz, Nina Seemann, Fabienne Braune, Andreas Maletti
- *Dynamic Models in Moses for Online Adaptation* by Nicola Bertoldi
- *Integrating a Discriminative Classifier into Phrase-based and Hierarchical Decoding* by Aleš Tamchyna, Fabienne Braune, Alexander Fraser, Marine Carpuat, Hal Daumé III, Chris Quirk
- *Large-scale Human Evaluation of Machine Translation Output for WMT 2013* by Christian Federmann, Matt Post
- *Makefiles for Moses* by Ulrich Germann
- *morphogen: Translation into Morphologically Rich Languages with Synthetic Phrases* by Eva Schlinger, Victor Chahuneau, Chris Dyer
- *MTMonkey: A Scalable Infrastructure for a Machine Translation Web Service* by Aleš Tamchyna, Ondřej Dušek, Rudolf Rosa, Pavel Pecina
- *Open Machine Translation Core: An open API for Machine Translation systems* by Ian Johnson
- *Pipeline Creation Language for Machine Translation* by Ian Johnson
- *QuEst - A framework for machine translation quality estimation* by Kashif Shah, Eleftherios Avramidis, Ergun Biçici, Lucia Specia
- *RankEval: Open tool for evaluation of machine-learned ranking* by Eleftherios Avramidis

- *Updating the Feature Function Framework in the Moses Decoder* by Hieu Hoang, Kenneth Heafield, Barry Haddow, Matt Post, Eva Hasler, Phil Williams, Chris Dyer, Philipp Koehn

- *Visualization, Search and Analysis of Hierarchical Translation Equivalence in Machine Translation Data* by Gideon Maillette de Buy Wenniger, Khalil Sima'an

- *XenC: an open-source tool for data selection in Natural Language Processing* by Anthony Rousseau

Given the large number of accepted submissions, we opted to present all of them as posters, with a boaster session giving 3–5 minutes to each presenter. In our experience and also based on the feedback we received, this was a good idea.

After projects and project presentations, research papers are the third most important component of MT Marathon: 91% of respondents want to keep them.

3.4 Invited Talks

This year we had 5 invited talks, one on each day:

- *Moses at the European Commission*, Francis Tyers (Prompsit)

- *Compositional Semantics, Deep Learning, and Machine Translation*, Phil Blunsom (University of Oxford)

- *Machine Translation Challenges, Solutions, and Applications*, Bonnie Dorr (DARPA)

- *Morphological Knowledge in Statistical Machine Translation*, Kristina Toutanova (Microsoft Research)

- *Domain Adaptation Using Parallel and Comparable Corpora*, Alex Fraser (Ludwig-Maximilians-Universität München)

DARPA and Microsoft Research sponsored the travel costs of their employees and MosesCore covered the remaining invited speakers.

Each of the invited talks was very attractive in its way, providing unique experience or visions. The comments in our feedback form indicate that the technical level of detail was just right: people complained about both too much linguistics and too little linguistics in the talks. One participant

suggested to somehow merge invited talks into the morning lectures where appropriate, since (some of them) provide a great introduction to emerging research. This may be an opportunity for some compression of the programme.

In our survey, none of the respondents said they would have deliberately skipped the talks. On the contrary, 49% of the respondents fully followed the talks and another 20% not only fully followed them but would have liked more.

3.5 "Summer" School

The summer school is a series of lectures with accompanying labs designed to provide a full introduction to statistical MT.

3.5.1 Lectures

The following is a list of the lectures in the summer school this year:

- *Introduction to MT*, Chris Dyer (CMU)
- *MT Evaluation and Quality Estimation*, Lucia Specia (Sheffield)
- *Word-Based Models (Word-Alignment)*, Adam Lopez (JHU)
- *Phrase-Based Models (Phrase Extraction)*, Mark Fishel (Zurich)
- *Decoding for Phrase-based Models*, Alexandra Birch (UEDIN)
- *Language Modelling (Theory, Practice)*, Marcello Federico (FBK) and Kenneth Heafield (UEDIN/CMU)
- *Hierarchical Models and Chart Decoding*, Barry Haddow (UEDIN)
- *Constituency vs. Dependency, TectoMT: Deep Syntactic Transfer, Error Correction*, Ondřej Bojar, Martin Popel and Rudolf Rosa (ÚFAL)
- *Discriminative Training*, Phil Blunsom (Oxford)
- *Computer Aided Translation and Integration with MT*, Marcello Federico (FBK)

This year, MosesCore supported travel expenses of four invited lecturers (Chris Dyer, Phil Blunsom, Adam Lopez and Lucia Specia).

Based on the feedback form, lectures were well attended, 49% of respondents paid full attention to them (and some of those would have even like

more lectures). Some respondents found the lectures too introductory or would prefer some more recent models, some found the level of detail and timeliness just right and enjoyed how the lectures build on one another during the week. Since 87% of respondents said they like the lectures as they are now, we probably won't make an error if we keep the structure and (as we have already been doing) update the topics only a little bit, based on the availability of lecturers.

3.5.2 Labs

This year we had five labs. The first one, on Monday, was an introduction to Moses and an experiment management system (Eman) used by the local organizer. The second lab was a repetition from last year's Tuesday/Wednesday mini-project, introducing people to word alignment algorithms.

There were two labs on Thursday, each devoted to one complex framework for more (though not exclusively) rule-based MT, namely Treex and Apertium.

The Friday lab was targeted at Moses developers: Hieu Hoang presented the proper way of adding a new feature to the model in the refactored Moses code.

Based on the feedback form, only 18% of respondents were fully involved in the labs. That is probably not surprising, since different labs targeted different groups of people from a rather wide spectrum of experience and background. Again, 84% of respondents suggest to keep the (structure of the) labs as they are. Opinions vary on the utility of labs: some (including us) expect that the labs are very good for novices, some would prefer people to work on projects instead, some liked the competitive lab on alignment best, some said it consisted mainly of homework assignment... Since all the labs are an optional part of the programme, it is probably alright to preserve the current structure.

3.6 Round table on large data, cluster setups, best parallelization practice

To advance the exchange of experience and sharing of best practice, MTM13 included a round table or rather a panel on computing resources and cluster configurations. The 90-minute long discussion started with seven short presentations given by:

- Alex Fraser (Uni. Stuttgart and LMU-Munich)
- Nicola Bertoldi (FBK)

- Barry Haddow (UEDIN)
- Matt Post (JHU)
- Matthias Huck (RWTH Aachen University)
- Ondřej Bojar (CUNI)
- Fethi Bourages (LIUM)

Each speaker gave a brief description of the computer setup at their institute and some presented additional useful (custom) tools, e.g. a web-based load monitor used at UEDIN or a generic tool for caching model files on local disks to relieve the disk server developed at RWTH Aachen University.

A short discussion followed or interleaved each of the presentations, confirming that the level of technical detail was just right for the audience.

3.7 Tutorial: Moses on EC2

For those that do not have access to a department or university cluster, Chris Dyer (CMU) presented the Amazon Elastic Cluster (EC2) service. The session was a mix of an introductory lecture of the cluster setup, and self-paced attempts to get the service running.

Chris Dyer negotiated sponsorship of the session from Amazon, so all participants were given 100 USD vouchers for Amazon computing services. The main message learnt was that computation power has indeed become a commodity and that it is relatively easy to set up the environment to run Moses and other MT tools, even in parallel on a custom SGE cluster.

3.8 TAUS Moses Round Table

On the third day of the Machine Translation Marathon in Prague, TAUS conducted a round table for commercial users. A total of 26+ people attended (there were about 5-10 unregistered walk-ins). The round table had broad participation from industry participants, government, representatives of the MosesCore team (Philipp Koehn, Barry Haddow and Hieu Hoang from University of Edinburgh and Ondřej Bojar from Charles University in Prague) and some researchers interested in the topic.

Rahzeb Choudhury (TAUS) set the theme of the round table introducing current and potential future areas of cooperation of Moses users from the industry: sharing knowledge, sharing investment and sharing code.

Each of the attendees then introduced themselves and briefly explained their use of Moses and what improvements they would like to see.

Achim Ruopp (TAUS) presented the results of the 2013 Moses User Survey and an analysis of the project progress from an industry perspective over the last 3 years (the detailed report is available for download for registered TAUS users[4] or without registration [5]). He compared Moses to similar open source projects with academic origins and what makes them successful or not in the commercial space.

Philipp Koehn then presented the latest developments of the Moses project and future plans. He also explained for the commercial users the way the Moses project works as an academic project.

One of the participants asked if a live demo of a Moses system (as a UI or web demo) is available and TAUS showed the Windows UI developed by Capita as part of the MosesCore project – the developer Jie Jiang was present and provided background. For information on MT system training TAUS showed the online tutorial and mentioned the step-by-step tutorial in the Moses documentation.

Initial discussion before break

The discussion then focused on the diagram in the report showing the current separation between the open Moses components and proprietary components built on top of them. Like in the report there were differing opinions where this line should be drawn in the future. The attendees discussed how the current open components should be developed and released in order to enable easy development of industry solutions like a UI or CMS integration.

Some attendees thought that the community is about two-thirds of the way towards an entirely open solution, including the easy integration into open source TMS/CAT tools. One attendee mentioned that information for easy install (and use) should be made more discoverable in the documentation. Francis Tyers (Prompsit) repeated the request for additional how-to documentation he already mentioned in his earlier MTM presentation "Moses at the European Commission".

2nd half

After the break Rahzeb Choudhury aligned the most requested improvements with different adoption phases/user groups and the round table attendees discussed these areas in depth:

Installing and using Moses – Beginners The consensus was that the resources available (Moses site, support list, MT and Moses Tutorial) are

[4] `https://www.taus.net/reports/are-moses-users-seeking-common-ground`

[5] `http://www.statmt.org/mosescore/index.php?n=Main.Publications`

sufficient, however it needs to be ensured that beginner resources are easily discoverable and that the documentation stays up-to-date and is easy to use. The v1 release is very welcome and the industry looks forward to future releases.

Integrating Moses into Existing Workflows/Systems – Implementation The main areas of cooperation (APIs and formatting) on integration are covered by current activity. TAUS will help with the next steps for Moses4Loc (Formatting) to help ensure there is thorough testing and adoption by the industry.

In this context it was mentioned that the MateCat project aims to provide XLIFF support and Andrzej Zydroń from XTM-INTL mentioned that supporting TIPP (TMS Interoperability Protocol Package) would be good.

Hieu Hoang urged to not duplicate efforts in tag handling and to check in the code into the Moses repository.

Training and Translation Speed – Production One idea in this area was providing guidance in the documentation how to use the latest speed enhancements for training and translation. Participants also would like guidance on required minimal system configurations. The Moses team pointed out that this depends very much on the amount of data used to train an MT system; however it might be possible to outline some typical hardware configurations.

Hieu Hoang asked about forums where to find out about requirements by the industry. RT participants mentioned Localization World, LinkedIn and GALA.

Data and Sharing Engines TAUS showed TAUS Data as a possible source for training data for SMT engines and described the TAUS Developing Talent initiative. Philipp Koehn mentioned that for academic papers it is required that other researchers can reproduce the results stated in a paper, for which freely available data might be necessary. TAUS is looking into the release of some older data.

The last area discussed during the round table concerned the sharing of trained engines. The Moses team pointed out that currently a trained engine depends on a whole chain of tools and training parameters, including for example tokenization tools and tuning data, so to produce the best results, the use of a trained engine requires the use of the same tool chain for translation with the engine.

The trained engine itself currently contains no meta-data about the used training environment. Ensuring a tool match for translation is currently the responsibility of the user. Defined releases are definitely making this a little easier.

Apart from this interoperability issue the other obstacles to sharing engines mentioned where potential copyright issues and privacy issues. An attendee mentioned that the availability of open-source anonymisation tools would be desirable.

3.8.1 Feedback on TAUS Round Table

The responses in our feedback form indicate that 53% of MT Marathon participants either attended TAUS Round Table or wanted to but in the end did not make it. This is a good result for this industry-oriented activity in an event that originates in academia. It is also noteworhy that 16% of all respondents would have liked even more of this.

4 Assessment

Based on the positive feedback from the participants, we are confident that MT Marathon 2013 was a successful event. The attendance of the whole event and all its parts was better than expected and the programme was broad enough to provide something for everyone at all levels.

The format of the event has been more or less stable throughout the years and this makes the tradition of MT Marathons stronger. Participants know what to expect and word of mouth spreads awareness about the event among students and – we hope – also users from the industry. The mix of introductory lectures and labs with advanced invited talks and research papers, and most importantly the group projects make the programme attractive both for newcomers as well as regular attenders. The stimulating environment provably allows new students to jump-start their research career in machine translation or natural language processing in general. With more industry-oriented projects (and we have already seen one this year), MT Marathons could even facilitate smooth transitions from study or research to commerce.

We are delighted to have received feedback from almost half of the participants. We specifically questioned whether the quite packed programme and structure of MT Marathons is accepted well or whether the participants would prefer some changes. The conclusion is that, indeed, the programme is perceived as large in content but that none of the marathon activities should

be abandoned. The projects are in general seen as the most important aspect and deserve some more space (and about the same level of organization, i.e. boaster session and public reporting). The remarks in the questionnaire indicate that it is probably the "extra" activities that should be cut first, i.e. no additional round tables (except the industry-oriented TAUS round table). Some further savings can be obtained by (partially) merging introductory lectures and invited talks, although people really liked the quality and quantity of invited talks this year.

A few specific issues were mentioned in the comments:

- The registration form needs to send out confirmation e-mails with the details. This was an omission by the local organizer and should be avoided in the future.
- Unlike last year, people did not complain about the programme being announced too late (the preliminary programme was first announced in late July). On the other hand, some participants were worried that they can't track changes in the programme and that the label "Preliminary" was removed only on the first day of the event.
- A little bit more structure for social events is perhaps desirable. Restaurant suggestions would be valuable, one participant suggested to include a guided tour of the city and Trento, the venue of the next MT Marathon, asks for an organized mountain hike.

Overall, we were very happy with this MT Marathon and look forward to the next one.

A Feedback Form

The following pages contain the printed version of an online feedback form sent to all participants of MT Marathon 2013.

MT Marathon 2013 - Feedback Form

* Required

1. **Where do you come from?** *
   What best describes your current occupation?
   Check all that apply.

   - [ ] undergrad (studying for Master)
   - [ ] postgrad (studying for Ph.D.)
   - [ ] postdoc (Ph.D. finished, young researcher)
   - [ ] researcher in a research institute or university
   - [ ] (small) academic research group leader
   - [ ] translator in a company / freelancer
   - [ ] researcher in a company
   - [ ] developer in a company
   - [ ] manager in a company

2. **How do you feel about the following toolkits after the MT Marathon?**
   Mark only one oval per row.

	Knew well enough before	Confident I can use it on my own	Not afraid, but will seek assistance	Still afraid	Didn't use this toolkit during this Marathon
Moses	◯	◯	◯	◯	◯
Joshua	◯	◯	◯	◯	◯
cdec	◯	◯	◯	◯	◯
TectoMT/Treex	◯	◯	◯	◯	◯
Apertium	◯	◯	◯	◯	◯

3. **Any other tools or useful toolkits you learned about?**
   Provide names or even links to toolkits that you did not know before at all or were not familiar with them and this MT Marathon allowed you to start using them.

   .....

   .....

   .....

   .....

   .....

4. **How much did you attend to various regular MTM tracks?** *

(sitting in the lecture theatre but working on a project counts as project work ;-)
Mark only one oval per row.

	Did not attend (Did not want to)	Did not attend (But wanted to)	Intermittently	Fully involved	Fully involved and would have liked more
Introductory morning lectures					
Keynote talks					
Labs					
Poster and demo presentations					
Work on projects					
Following other project reports					
General networking (meeting people)					

5. **...and to extras this year?** *
This Marathon featured irregular sessions that we may or may not repeat or modify in the coming years, e.g. run a panel discussion on a different topic. How much did you follow these tracks this year?
Mark only one oval per row.

	Did not attend (Did not want to)	Did not attend (But wanted to)	Intermittently	Fully involved	Fully involved and would have liked more
TAUS Round Table for Commercial Users					
Round Table on Clusters					
Tutorial on EC2					

6. **Is there anything you would suggest changing in next Marathons?** *
Mark only one oval per row.

	Keep	Change (details below)	Drop altogether	I don't care
Introductory lectures				
Keynote talks				

	Keep	Change (details below)	Drop altogether	I don't care
Labs	◯	◯	◯	◯
Papers on tools (presented as posters and demos this year)	◯	◯	◯	◯
Projects	◯	◯	◯	◯
Project presentations	◯	◯	◯	◯

Detailed comments

Please tell us what did you like or not like about each of the activities. How could they be improved. (This is the place to propose any changes.)

7. **Introductory lectures**

.....

.....

.....

.....

.....

8. **Keynote talks**

.....

.....

.....

.....

.....

9. **Labs**

.....

.....

.....

.....

.....

10. **Paper presentations (posters this year)**

.....

.....

.....

.....

.....

11. **Projects**

.....

.....

.....

.....

.....

12. **Project presentations**

.....

.....

.....

.....

.....

13. **Project follow-up reports**
    Some Marathon projects will run longer. We have no control over what is going to happen with them, but still: is there anything specific we should try to make you benefit more from such on-going projects?

.....

.....

.....

.....

.....

14. **Extra activities**
    Here is the place to comment our extras (round tables etc.). Is there a particular topic or activity you would like to take part in in the coming years?

.....

.....

.....

.....

.....

15. **How should we present the tool papers?** *
    Mark only one oval.

- ◯ Talks (as before 2013)
- ◯ A mix of talks and posters
- ◯ All posters (as in 2013)
- ◯ I don't care

16. **Was the information before Marathon sufficient?**
    We tried to provide you with all relevant information on the web page early just in time to help you in all decisions and steps. Was there anything missing? Would you like to have received more details, at different times or in a different manner?

    .....

    .....

    .....

    .....

    .....

17. **Anything you want to add? Any other comments?**
    Any other impact or impression MT Marathon 2013 has made on you?

    .....

    .....

    .....

    .....

    .....

B Automatic Summary of Responses

The following pages contain the printed version of a detailed automatic summary of all the responses we collected using our online form.

ondrej.bojar@gmail.com

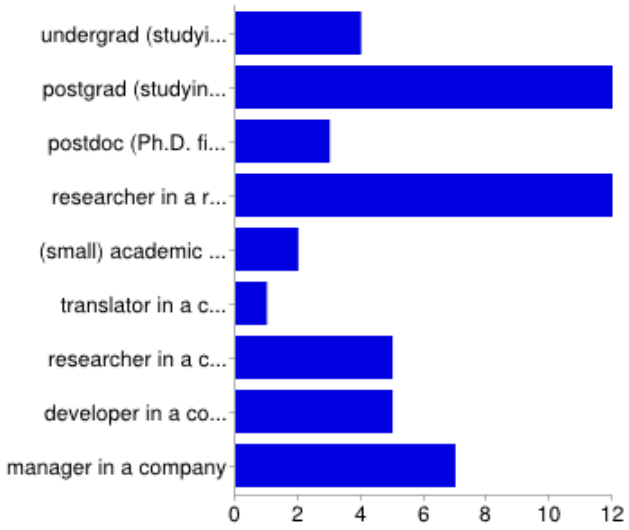
Edit this form

45 responses

View all responses

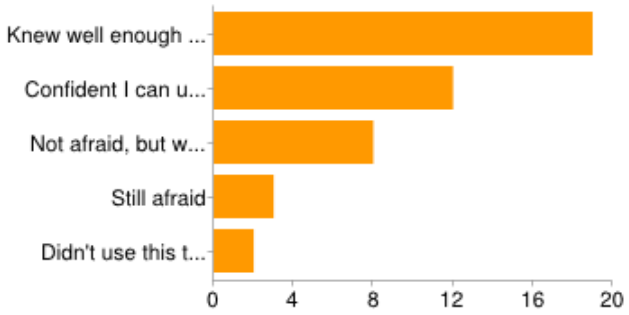
Summary

Where do you come from?



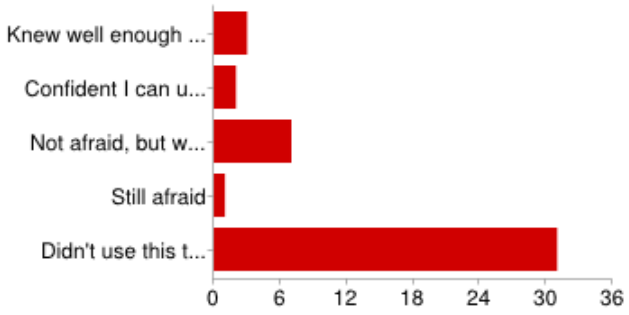
undergrad (studying for Master)	**4**	8%
postgrad (studying for Ph.D.)	**12**	24%
postdoc (Ph.D. finished, young researcher)	**3**	6%
researcher in a research institute or university	**12**	24%
(small) academic research group leader	**2**	4%
translator in a company / freelancer	**1**	2%
researcher in a company	**5**	10%
developer in a company	**5**	10%
manager in a company	**7**	14%

Moses [How do you feel about the following toolkits after the MT Marathon?]



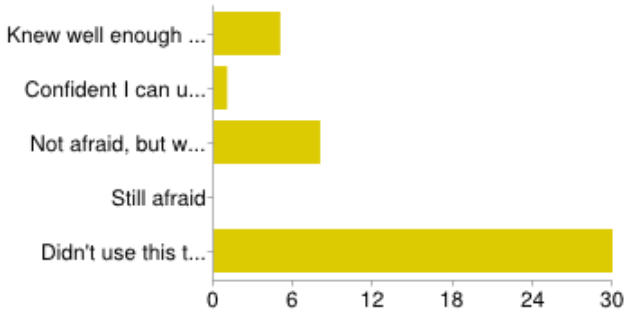
Knew well enough before	**19**	43%
Confident I can use it on my own	**12**	27%
Not afraid, but will seek assistance	**8**	18%
Still afraid	**3**	7%
Didn't use this toolkit during this Marathon	**2**	5%

Joshua [How do you feel about the following toolkits after the MT Marathon?]



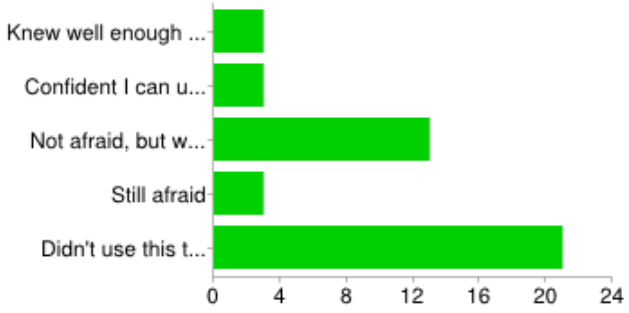
Knew well enough before	**3**	7%
Confident I can use it on my own	**2**	5%
Not afraid, but will seek assistance	**7**	16%
Still afraid	**1**	2%
Didn't use this toolkit during this Marathon	**31**	70%

cdec [How do you feel about the following toolkits after the MT Marathon?]



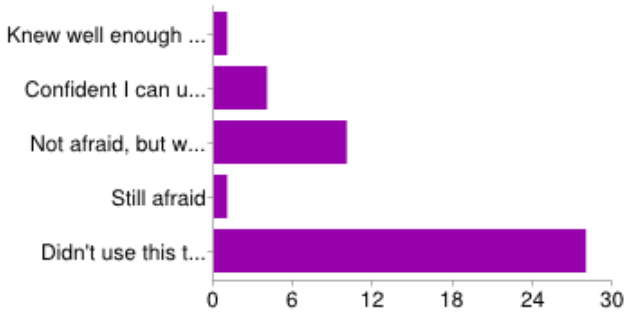
Knew well enough before	**5**	11%
Confident I can use it on my own	**1**	2%
Not afraid, but will seek assistance	**8**	18%
Still afraid	**0**	0%
Didn't use this toolkit during this Marathon	**30**	68%

TectoMT/Treex [How do you feel about the following toolkits after the MT Marathon?]



Knew well enough before	**3**	7%
Confident I can use it on my own	**3**	7%
Not afraid, but will seek assistance	**13**	30%
Still afraid	**3**	7%
Didn't use this toolkit during this Marathon	**21**	49%

Apertium [How do you feel about the following toolkits after the MT Marathon?]

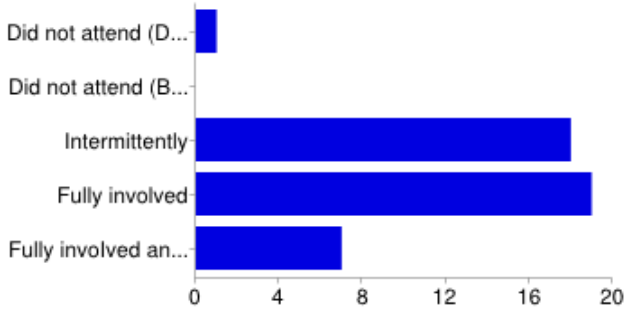


Knew well enough before	**1**	2%
Confident I can use it on my own	**4**	9%
Not afraid, but will seek assistance	**10**	23%
Still afraid	**1**	2%
Didn't use this toolkit during this Marathon	**28**	64%

Any other tools or useful toolkits you learned about?

ASIA ON LINE QuEst for Machine translation Vowpow Wabbit yes I foun lots of useful new things in the poster session. Most tools presented seems useful, I'm testing some of them right now. I heard about the MT Kitchen, but did not really learned the details of how to use it I learned about VowPal Wabbit, for discriminative classification http://hunch.net/~vw/ Supposedly very handy when you want to do phrase sense disambiguation etc for domain adaptation starcluster, EC2 I learned about Casmacat and MateCat, but not enough to start using them. Eman ! Quest, RankEval KenLM Eman eman KenLM QuEst

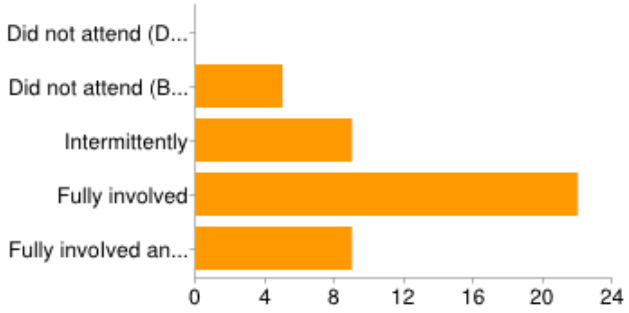
Introductory morning lectures [How much did you attend to various regular MTM tracks?]



Did not attend (Did not want to)	**1**	2%
Did not attend (But wanted to)	**0**	0%

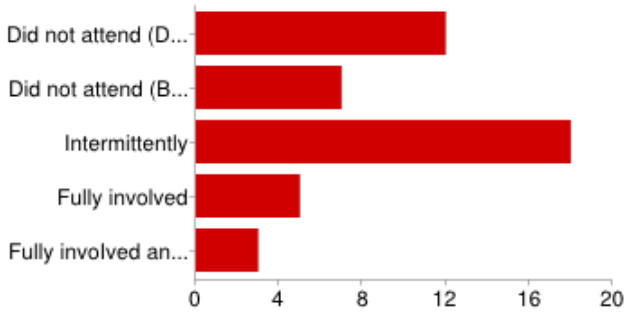
Intermittently	**18**	40%
Fully involved	**19**	42%
Fully involved and would have liked more	**7**	16%

Keynote talks [How much did you attend to various regular MTM tracks?]



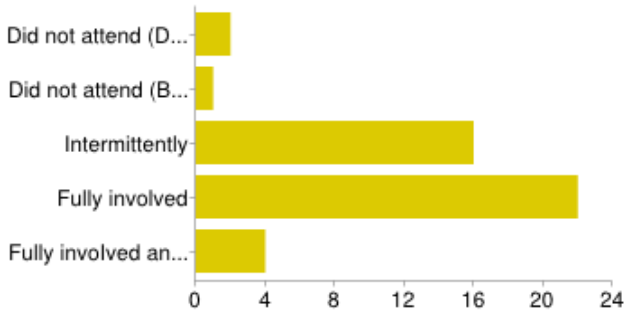
Did not attend (Did not want to)	**0**	0%
Did not attend (But wanted to)	**5**	11%
Intermittently	**9**	20%
Fully involved	**22**	49%
Fully involved and would have liked more	**9**	20%

Labs [How much did you attend to various regular MTM tracks?]



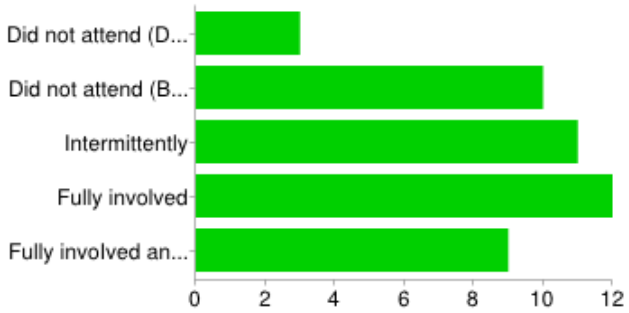
Did not attend (Did not want to)	**12**	27%
Did not attend (But wanted to)	**7**	16%
Intermittently	**18**	40%
Fully involved	**5**	11%
Fully involved and would have liked more	**3**	7%

Poster and demo presentations [How much did you attend to various regular MTM tracks?]



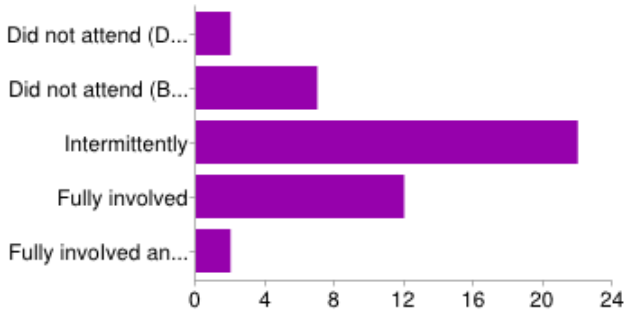
Did not attend (Did not want to)	**2**	4%
Did not attend (But wanted to)	**1**	2%
Intermittently	**16**	36%
Fully involved	**22**	49%
Fully involved and would have liked more	**4**	9%

Work on projects [How much did you attend to various regular MTM tracks?]



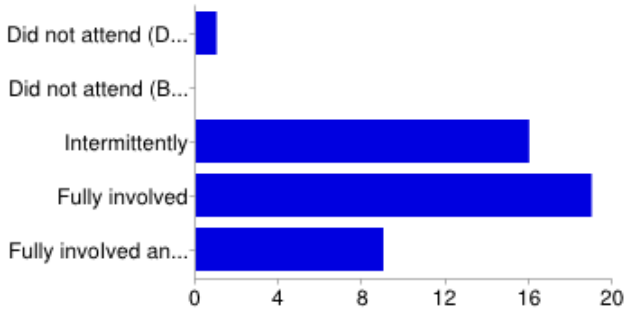
Did not attend (Did not want to)	**3**	7%
Did not attend (But wanted to)	**10**	22%
Intermittently	**11**	24%
Fully involved	**12**	27%
Fully involved and would have liked more	**9**	20%

Following other project reports [How much did you attend to various regular MTM tracks?]



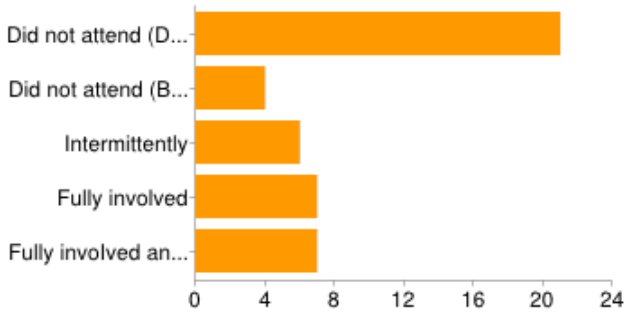
Did not attend (Did not want to)	**2**	4%
Did not attend (But wanted to)	**7**	16%
Intermittently	**22**	49%
Fully involved	**12**	27%
Fully involved and would have liked more	**2**	4%

General networking (meeting people) [How much did you attend to various regular MTM tracks?]



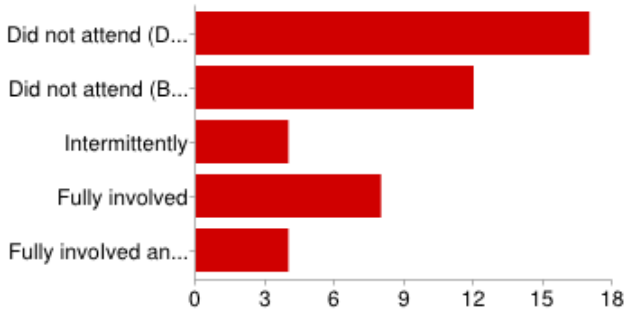
Did not attend (Did not want to)	**1**	2%
Did not attend (But wanted to)	**0**	0%
Intermittently	**16**	36%
Fully involved	**19**	42%
Fully involved and would have liked more	**9**	20%

TAUS Round Table for Commercial Users [...and to extras this year?]



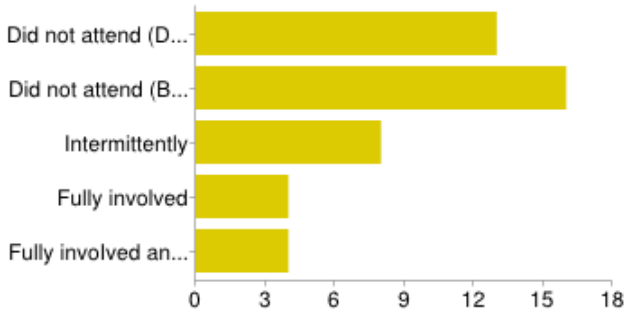
Did not attend (Did not want to)	**21**	47%
Did not attend (But wanted to)	**4**	9%
Intermittently	**6**	13%
Fully involved	**7**	16%
Fully involved and would have liked more	**7**	16%

Round Table on Clusters [...and to extras this year?]



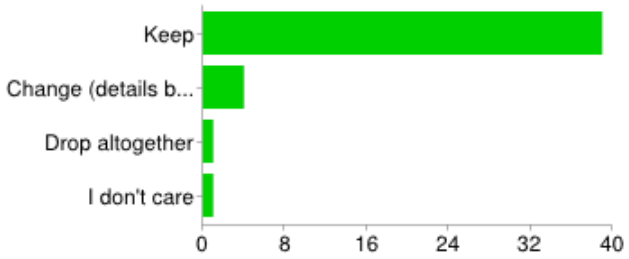
Did not attend (Did not want to)	**17**	38%
Did not attend (But wanted to)	**12**	27%
Intermittently	**4**	9%
Fully involved	**8**	18%
Fully involved and would have liked more	**4**	9%

Tutorial on EC2 [...and to extras this year?]



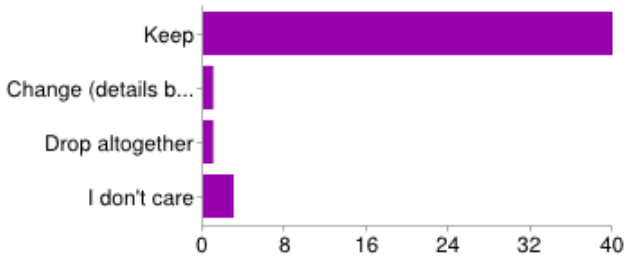
Did not attend (Did not want to)	**13**	29%
Did not attend (But wanted to)	**16**	36%
Intermittently	**8**	18%
Fully involved	**4**	9%
Fully involved and would have liked more	**4**	9%

Introductory lectures [Is there anything you would suggest changing in next Marathons?]



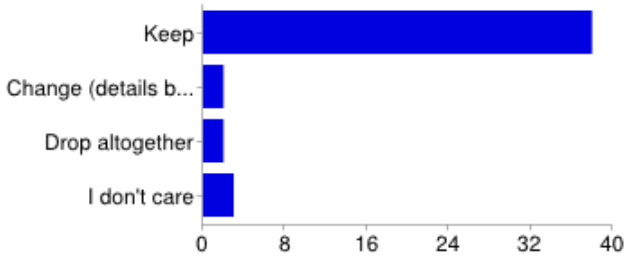
Keep	**39**	87%
Change (details below)	**4**	9%
Drop altogether	**1**	2%
I don't care	**1**	2%

Keynote talks [Is there anything you would suggest changing in next Marathons?]



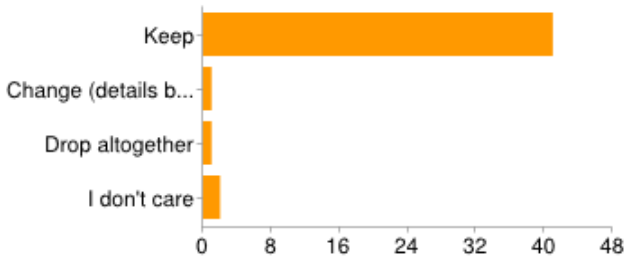
Keep	**40**	89%
Change (details below)	**1**	2%
Drop altogether	**1**	2%
I don't care	**3**	7%

Labs [Is there anything you would suggest changing in next Marathons?]



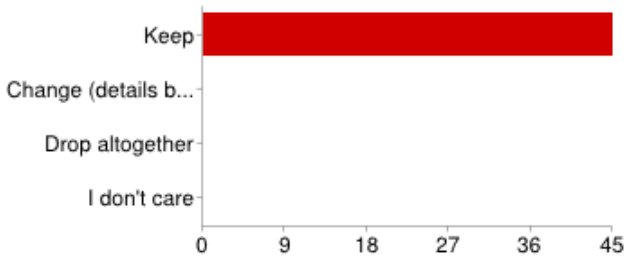
Keep	**38**	84%
Change (details below)	**2**	4%
Drop altogether	**2**	4%
I don't care	**3**	7%

Papers on tools (presented as posters and demos this year) [Is there anything you would suggest changing in next Marathons?]



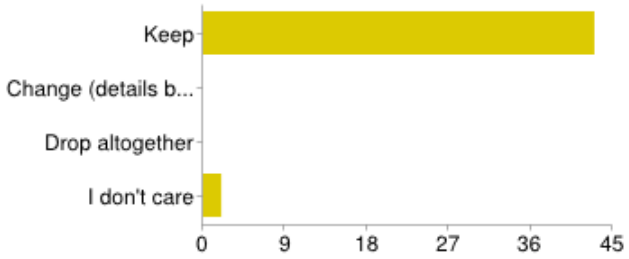
Keep	**41**	91%
Change (details below)	**1**	2%
Drop altogether	**1**	2%
I don't care	**2**	4%

Projects [Is there anything you would suggest changing in next Marathons?]



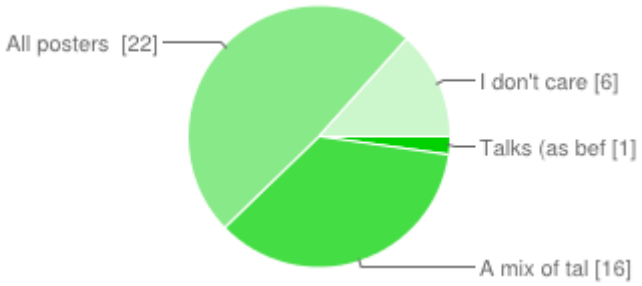
Keep	**45**	100%
Change (details below)	**0**	0%
Drop altogether	**0**	0%
I don't care	**0**	0%

Project presentations [Is there anything you would suggest changing in next Marathons?]



Keep	**43**	96%
Change (details below)	**0**	0%
Drop altogether	**0**	0%
I don't care	**2**	4%

How should we present the tool papers?



Talks (as before 2013)	**1**	2%
A mix of talks and posters	**16**	36%
All posters (as in 2013)	**22**	49%
I don't care	**6**	13%

Number of daily responses

B.1 Detailed Textual Comments

B.1.1 Introductory lectures

- Some too introductory (this coming from someone who had only been involved in MT for a few months).
- Some of the introductory lectures like the one about decoding for phrase-based models could be more attractive. The topic is very interesting but the lecturer seemed not to be really into present the topic.
- I liked how you connected the lectures to each other, so they built on

one another. A minor historical detail: it was presented that the influence of the ALPAC report lasted maybe 25 years until statistical methods took over. I believe you might say rather 10-15 years, whereupon a lot of rule-based systems were developed. It might be useful at that point to bring up the problems researchers were facing with rule-based systems, that statistical methods helped out with. E.g. the 70-30 or 80-20 rule: it's easy to capture 70-80% of grammatical generalizations, but extremely hard for the final 20-30%. Huge, finely-tuned rulesets became delicate, difficult to maintain, and hard to adapt to language change and different domains.

- It seemed to me that for at least some of the lectures, many of the attendees were already established researchers. A few people who delivered lectures made the same observation. Perhaps the balance could be altered a little so that there was some introductory information and some "this is what other people are doing - this is what is state of the art" in each lecture.

- more practice

- Very good presentations. Clear and inspiring, even for people who have already a lot of experience in the field. These are good, I think. It provides some introductory material for students who are newer to MT.

- A single lecture is probably too short to learn about, say, phrase-based decoding and thus the benefit for the intended audience, supposedly undergrad is small. The time slots would be better used showing demos of tools.

- I loved them all, although I attended only a few.

- Basically everything was SMT - a broader coverage of approaches would be a lot more interesting to me.

- I would question how much need there is to continue with the very basic MT lectures (eg PB modelling and decoding). Maybe we could compress them to make for more advanced lectures on new topics. e.g. continuous space models.

- Very informative and well explained, no potential for improvement from my side whatsoever!

- The lectures were perfect, they gave the right overview of everything so that one gets the big picture and is ready to both start dealing with MT practically (also thanks to the labs) and to study more about MT from papers. But I would still like to hear more about some new state-of-the-art things currently taking place in MT; I feel that the lectures are more of talking about what SMT was like maybe a year or two ago... But maybe I am wrong.

B.1.2 Keynote talks

- Mostly good. Could have been more technical (felt a little dumbed down?). Although maybe at the right level for people with only a linguistics background.
- I would have loved to listen to someone talking more about the linguistic aspects of MT
- Very high quality and good. Only the talk by Darpa was somewhat less interesting from a research point of view as it was somewhat too general, but it was not bad also.
- The two talks by Phill Blunsom were very clear, deep and well-rounded. Very inspiring as well.
- My personal favorite was 'Compositional Semantics, Deep Learning, and Machine Translation'. I would have liked to learn more about how variable-length language data is transformed into fixed-length vectors in ways that preserve information. I also liked the overviews from EU and DARPA.
- more innovative
- Recoding the talks is the right thing to do.
- Very informative and well explained, no potential for improvement from my side whatsoever!
- I think these should be done away with and/or folded into the introductory lectures. e.g., Kristina and Phil's talks were great introductions to emerging research in MT, so they could be part of the intro lectures. I guess Bonnie's talk was interesting, but (a) not really useful to students and (b) not new at all to senior researchers.

B.1.3 Labs

- I can only account for Adam's alignment lab and that was very good and sparked considerable interested among participants.
- Labs are too short to learn anything properly
- I'm not sure that these are that useful. I think it would be better to have students focus on participating in the projects. Already there is not enough time, and this consumes even more.
- Perhaps there were too many this year, reducing time for projects? Having said that, all seemed interesting and were not obligatory.
- more tutorials
- I did not attend those, but what I hear from people who are novice they do benefit quite a lot from those. So definitely keep them I'd say.
- some better than others
- Maybe it would be an idea to include at least one or two basic sessions for Moses beginners so they can learn the basic concepts of the system.
- It would be good to try to split labs and projects so people could do both.
- I really liked the hands-on nature of the labs. The alignment one was the least useful for me, consisting mostly of a homework assignment we really didn't need to use limited conference time for. I was blown away the tree graph interface of Treex–really impressive!

B.1.4 Paper presentations (posters this year)

- Posters are good but since many of these are tools we could also have more demos. Nice. I think this is good, but I would have liked presentations still more than posters. On the other hand there is only a limited amount of time, and I really liked the high quality and great quality of the many keynote talks this year. So I think as a whole this was probably an improvement given there is only limited time.
- Posters are the best format for this kind of paper. I would not change them back to oral presentations. The boaster session was great.
- seemed ok

- Make sure that accepted papers are high-quality, and that the presented tools are relevant and will be used.

- Very informative and well explained, no potential for improvement from my side whatsoever! Seemed to work well as posters, at least as a presenter who didn't get a chance to walk around and look at other papers

- Posters is probably the best way of presenting tools. People can talk to get the answers they want from the author(s). Perhaps for some tools a presentation would be good since a demonstration can be done for a larger audience. A poster session can then be used for any deeper questions.

B.1.5 Projects

- Projects are the reason for the Marathon – they're where things get made, collaborations form, and research happens. There are more and more distractions every year. While it's reasonable to have lectures or tutorials in the morning, afternoons should be unencumbered by distractions.

- I missed the projects and I have made a mistake. I would have liked to be involved more - probably the projects occurred too late during the day

- One person commented that ironically since this year there were so many good talks, there was less time to work on the projects. This has some truth to it, but I still thinks many good talks is a very good thing. Also I really liked the opportunity the Projects give to work directly with people, and exchange ideas in a very informal way. As it is I would perhaps not have been able to implement soft-constraint decoding in Joshua without this MT Marathon!

- Given that most projects already have a presence on open source hosting sites (github, Google project hosting, ...) or it can be easily obtained, it is not necessary to set up a separate source control (svn); all project info should be pooled on the hosting site, so that projects can be easily continued after MTM; provide more info on available machine resources at the MTM location on info sheet (I spent extra time trying to use the student machines and ran out of disk space due to quota), cluster seemed to be complicated to use - make cloud resources available during MTM? (e.g. Amazon was sponsor?)

- Good selection of projects and interesting coverage of topics!
- To me, this is what MTM should be about, and it seems that every year, there is less and less time for this. I understand the temptation to add more material, but I think MTM should decide what it's main purpose is, and then to remove things that don't serve that purpose. As Steve Jobs said, the hardest part of running a company is saying "no" to things. I think MTM is saying "yes" to too many events (round tables, all kinds of lectures, etc) and diluting it's most useful function (working as a hackathon for people to get projects done).
- Because I went to all the lectures and labs, and was a little jet-lagged, I didn't get involved in a project, which was a mistake, I think.

B.1.6 Project presentations

- Enjoyable and good to have some final result for the week and share it with everyone. Also good as documentation to see later what everyone was working on when you want to get again in contact with people.
- maybe a bit too formal, since not everyone had slides to present during the updates
- I liked the way they were presented
- It is good to have a wrap-up presentation to see how far everyone got. Definitely keep this.
- Very well done and interesting!

B.1.7 Project follow-up reports

- See above - host projects on publicly available open source sites like github
- Perhaps ask people to update their slides on the wiki once the project is done and notify the mailing list with all participants when this is done?
- It really depends on the participants not on the MTM organisers.
- I don't know. I think it depends mainly on the interest of the team members to continue collaborating and sharing whether this will happen. In my case, the cooperation with JHU and the Joshua team that resulted from the project was very rewarding in any case.

- Very well done and interesting!

B.1.8 Extra activities

- For the next MT marathon in Trento I would not mind some extra social activities, if any time can be found in the Schedule. An afternoon of hiking in the mountains perhaps on the Sunday before ore on Saturday would certainly be very nice, just to give some suggestion!
- more on use of MT in industry
- The TAUS Roundtable was definitely interesting and I will attend it again!
- Probably some activities could in parallel. There is likely little overlap between audience of labs and industrial roundtables.
- There were too many of them. Cut back on them, as hard as it might be. Senior researchers spend all their time at such things and it detracts from progress on other projects.

B.1.9 Was the information before Marathon sufficient?

- Perfect
- It was very well organized!
- I did not get any confirmation of my registration, which made me wonder whether I am registered or not...
- Yes. Good communication, I was very satisfied with this
- It would be nice to receive a confirmation email when you book the accommodation. I did not receive any when I booked mine so I had to ask the secretary about what I had included in my booking (days, lunches, etc).
- Everything was great, wonderful organization!
- Great info! Preliminary program was sent out in late July, it was hard to track since then if the program had changed/what had changed. The program said "preliminary" short until the begin of the MTM.
- Yes, I think so. Yes, very helpful.

- yes
- Everything was clear.
- Yes, definitely.
- everything at time and enough
- Yes
- Sufficient.
- We should organize restaurant suggestions. It's good that we have the Saturday as a final day and not Friday.
- Website was good, but showed "preliminary" right up to the conference.
- The information was sufficient Yes.
- Everything was fine.
- Not that I noticed.

B.1.10 Anything you want to add? Any other comments?

- Great organisation! Picking the project on the blackboard was a good idea, transparent and quick. Please ask the caterers to leave food and especially water around. People get thirsty doing the projects!
- I enjoyed meeting SMT researchers, whom I did not know well. Overall everyone was very helpful and friendly, but sometimes difficult to reach for people coming from a RBMT background - probably a normal reaction/behavior.
- For me, the only bad thing about this year's MT Marathon was that I had to leave it much too early (Wednesday afternoon) because of private priorities - although I'd have enjoyed it the whole week long :-) A lot of things were rather new to me. But now I have a better overview on the subject of MT and hopefully can deepen it until next year. Last but not least - thank you for everything!
- Organization this year was outstanding! Thanks to the team in Prague! Also: The marathon is my favourite MT event in the year and Prague is my favourite place for a conference.

- Moses has a very active mailing list, but there are so many questions and answers that is very hard to keep track. It would be better to move to a kind of stack-overflow system were the questions and answers are online searchable, and somehow structured so that it is easier to find stuff and avoid repeated questions. Similarly, it would be good if the the main tools publish some documentation in the form of papers etc showcasing the main new functionality. I think the Joshua 2.0, 3.0 etc papers show one reasonably approach how this can be done, while benefiting the authors in the form of publications as well. One remark about the tool papers: I do actually care, but there was no category "other" Namely, I suggest either posters - if there are lots of high quality key note talks like in 2013 - or talks if there is more space in the program, due to less keynote talks. However, a combination of talks and posters seems undesirable, as it would be somewhat arbitrary and thus unfair who gets to give a presentation and who can only present a poster.
- <3 that, sure!
- Thanks!
- more tutorials
- Social activities: An allocated time slot with a guided tour through the city on one of the afternoons would have been a nice social event. Student involvement: I would very much appreciate the participation of many more grad and undergrad students. The organizers should actively advertise MT Marathon to young students who focus on NLP/CL. They would greatly benefit from the event. Introductory lectures and labs are useless if a huge majority of the registered participants are experts already.
- Ondrej & team organized it excellently! I believe this helped all participants to get the most out of their time at MTM.
- No
- I personally haven't installed Moses before, and heard about difficulty with setting it up and configuring it. A how-to-install-a-base-system lab on your own computer, and the options you have to proceed afterward, would have been useful to me personally. (I may be different from most there, though.
- Thanks for doing it!

- In general, the program was a bit too structured, which made it difficult to both attend things and work on projects, especially for someone who had to miss some things due to being sick on arrival in Prague.
- T H A N K S
- The MT marathon is getting a bit broad. I like the keynote talks, but I they are of substantially varying quality and the audience is very mixed. I think having a more narrow focus on the projects and the labs is something to consider.
- It was wonderful and very useful!

C List of Abbreviations

This report uses the following abbreviations:

API	Application Programming Interface
CAT	Computer Assisted Translation
CMS	Content Management System
CMU	Carnegie Mellon University
CSLM	Continuous Space Language Model
CUNI	Charles University in Prague
DARPA	Defense Advanced Research Projects Agency
FBK	Fondazione Bruno Kessler
GALA	Globalization and Localization Association (incl. the conference)
GHKM	Rule extraction algorithm by Galley, Hopkins, Knight, and Marcu (What's in a translation rule? In HLT-NAACL 2004).
JHU	Johns Hopkins University
LIUM	Laboratoire d'Informatique de l'Université du Maine
LMU	Ludwig-Maximilians-Universität München
MIRA	Margin Infused Relaxed Algorithm
MT	machine translation
MTM	MT Marathon
PBML	Prague Bulletin of Mathematical Linguistics
RWTH	RWTH Aachen University
SGE	(Sun/Son of) Grid Engine
SMT	statistical machine translation
SVN	Subversion, versioning software
TAUS	Translation Automaton User Society
TIPP	TMS Interoperability Protocol Package
TM	translation memory
UEDIN	University of Edinburgh
UI	user interface
WMT	Workshop on Statistical Machine Translation
XLIFF	XML Localization Interchange File Format