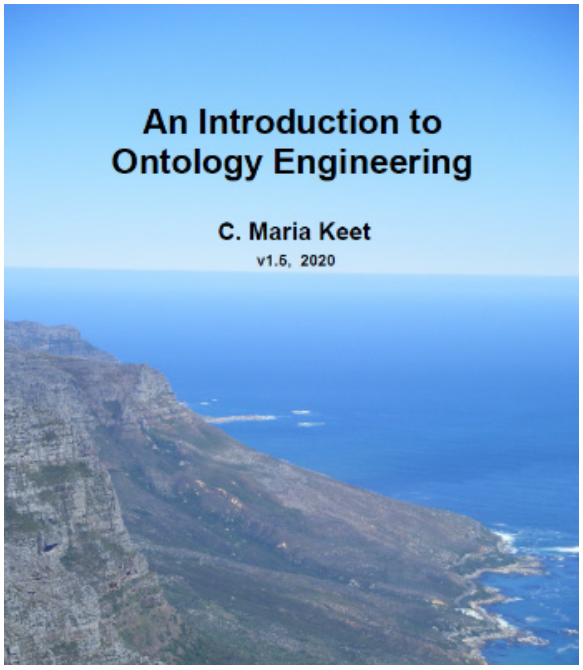
8

MARIA KEET'S OPEN TEXTBOOK JOURNEY

Grantee: Maria Keet
Position: Associate Professor
Department: Computer Science
Faculty: Science
Course: Ontology Engineering
Degree level: Postgraduate
Title of initiative: Ontology Engineering
Title of open textbook: *An Introduction to Ontology Engineering*

Introduction

Imagine if you could write the first textbook in your academic discipline. Would you produce an expensive, copyrighted, printed volume (the default choice) or would you avail yourself of new technological and legal innovations and produce a freely accessible, openly licensed, digital textbook (thereby establishing a powerful precedent)? This was the choice that faced Associate Professor Maria Keet in the Department of Computer Science at the University of Cape Town (UCT) who, in 2018, wrote the world's first textbook in the new subfield of ontological engineering, titled *An Introduction to Ontology Engineering*. The work emerged from her research interests in knowledge engineering with ontologies, concept modelling and related natural language generation, and from the fact that she was teaching an Ontology Engineering course in Computer Science which she felt would benefit from a dedicated textbook.



The Digital Open Textbooks for Development (DOT4D) project in the Centre for Innovation in Learning and Teaching at the University of Cape Town (UCT) provided grant funding and implementation support to 10 open textbook projects in the period from March 2018 – March 2019, as well as implementation support to an 11th initiative. The Open Textbook Journeys series tells the stories of the people driving these initiatives, their teaching and publishing processes and what inspires them to do this work. These case studies were developed in collaboration with and reviewed by the open textbook authors profiled.

Maria has an interest in open approaches towards knowledge production and dissemination and thinks that knowledge should be available free of charge so that it can increase possibilities for new research, scholarship and application. She feels that open educational resources play an important role in saving students (and interested lay people) money on expensive textbooks and that they can help redirect resources into parallel activities, such as content translation into different languages.

In February 2019, Maria was awarded a grant from the Digital Open Textbooks for Development (DOT4D) project to update and enhance her original *An Introduction to Ontology Engineering* work into a dynamic, interactive online textbook[1] with more content, more educationally relevant software, exercises and tutorials than found in the static PDF. She also increased the amount of locally relevant examples and content.

Maria's open textbook journey is fascinating because she has written the first textbook in her subfield, creating a compelling precedent for other educators to follow.

This case study tells her story, drawing on:

- Maria's grant proposal to the DOT4D project.
- Maria's grant report to the DOT4D project.
- Field notes of the DOT4D Publishing and Implementation Manager.

What is the problem Maria is trying to address?

Maria aims to fill the gap in educational material for ontology engineering that exists internationally – not only for the development of ontologies, but also to enable computer scientists to devise new methods, techniques and tools to facilitate the development of better-quality ontologies. Aimed at Honours and early postgraduate students, her work identifies and demarcates ontology engineering; introduces the reader to its essential components; and provides explainers, summarised and digested versions of scientific papers, as well as exercises to interactively engage with the material. At the time of writing, it is still the only textbook in ontology engineering for computer scientists.

Maria also aims to address the issue of cost savings for students, estimating that an imported commercially published textbook of this nature, if it existed, would cost in the region of R1,500–R2,000 ($100–$140).

Maria also aims to address the issue of cost savings for students, estimating that an imported commercially published textbook of this nature, if it existed, would cost in the region of R1,500–R2,000 ($100–$140)

Disciplinary conventions regarding textbooks

Maria indicated that, in her department, the decision about which textbook(s) to use for a course, including whether to use a textbook or not, is made by whoever teaches the course. For her, the main considerations when choosing a textbook include the:

- Level and primary envisioned audience.
- Topics covered and the order of topics, including its pedagogy.
- Degree to which it contains ancillary materials, such as adaptable slides, answers to exercises, code snippets and assistive software.
- Popularity and reputation of its authors.
- Cost and licensing (preferably open).

Maria indicated that her department tends to choose free resources wherever possible, or otherwise utilise a low-cost option to keep the costs down for students.

Maria's open textbook journey

Original plan

At the beginning of her open textbook journey, Maria set out to develop her resource through exploring how she could address the fact that there had not been any kind of ontology engineering textbook available for computer science students up until the first version she published in 2018. The writing of this first version began in 2009 with informal blogs which Maria improved and extended over many years. This initial version was made available as an open textbook on various platforms, such as Maria's website, the OpenUCT repository, and the Open Textbook Library in July 2018. Maria did however feel that the resource required content updates and enhancements in order for it to become a dynamic online textbook that includes software, additional exercises beyond those included already, tutorials and more locally relevant examples.

[1] https://people.cs.uct.ac.za/~mkeet/OEbook/

> The writing of this initial version began in 2009 with informal blogs which Maria improved and extended over many years

Authorship approach

Maria worked as a 'solo author' in terms of the production of the online textbook but has forged collaborations with colleagues and students to address certain aspects of the materials development process. She worked with Zubeida Khan (a UCT PhD graduate and now Senior Researcher at the Council for Scientific and Industrial Research) on a new chapter for the updated 2020 release version. She also engaged with current and former students in the process of updating material, as well as colleagues who could assist with foreign language aspects of the work.

The students involved in the project gained significant experience in writing and refining the tools presented in the textbook and the tool developers in particular expressed pride in being acknowledged in the software page accompanying the textbook.

> Maria held a 'typo contest' in order to get students to spot errors in the textbook

The content development process and student involvement

There were various deliverables that were achieved within Maria's grant period. Firstly, Maria produced a new chapter on modularisation in collaboration with Dr Khan. Following this, she carried out significant revisions to the multilingual aspects of the textbook. This included extending Chapter 9, compiling the multilingual Controlled Natural Language (CNL) literature list, adapting the ontology-based software development environment (ODE) localiser in a reusable way for Afrikaans and Spanish, and including instructions for localisation in other languages.

In an effort to promote student involvement in the content development process, Maria held a 'typo contest' in order to get students to spot errors in the textbook.

In addition, more exercises and answers were created, with an increase of about 10% to the 114 exercises that were part of the originally published textbook. This process also included two new tutorials on OntoClean and Ontology-based Data Access (OBDA) with data on South African national parks and elephant populations. The OntoClean tutorial was incorporated as an appendix in the new version of the textbook with adjustment of authorship.

More software has been developed and improved in Maria's textbook production process, enabling better support for the exercises and transitions between chapters in the textbook. This is now also presented systematically for all relevant tools. The two main tools that were enhanced as part of the project are the Description Logic (DL) plugin and the ODE localiser. In addition, instructions for access to the DL symbols for the visually impaired, and writing with DL symbols, were written by colleague and UCT PhD graduate Joan Byamugisha and made available online.

Lastly, the slides within the initial textbook were updated, cleaned up, converted from LaTeX into PowerPoint, and both versions were made available on the textbook page.

These last four aspects of the content development process greatly contributed to turning the textbook from 'just a PDF' into the interactive textbook that was expected by students and lecturers. The new and localised content that was produced resulted in a 10% increase of the main text and a 30% increase in the number of appendices due to the addition of tutorials and answers to the exercises.

Several people were involved in the series of updates entailed in producing the online textbook, including current and former students of the course. Maria described how the students involved in the project gained experience in writing and refining the tools. In addition, the tool developers expressed pride in being included in the software page of the textbook and being recognised as contributors in one of Maria's blog posts. She stated that, 'this, perhaps, reflects a culture in computer science, where software design and development has a higher status than writing text (which is generally despised as an activity). While there is clearly a benefit to the students in learning to write in a teaching style, I doubt that was appreciated to that extent.'

Production and publishing

Maria adopted an 'author as self-publisher' approach to publication of her originally published 2018 textbook and her 2020 online textbook, in that she managed all the production and publishing aspects herself.

Maria indicated that she conducted extensive research on getting the 2018 edition published, but comparable books at all the publishers she checked ranged from R1,400–R2,500 ($96–$170), which was too expensive in her view. She also felt that an external publishing process would take too long and that the textbook was already overdue. After she had taken the self-publisher approach, a non-profit academic publisher, College Publications, took an interest in the work and published a hard copy version of it.[2]

Maria's decision to openly license her work under a Creative Commons licence helps to facilitate global reach, in that it enables adaptation and translation into a range of languages. Usage data from the OpenUCT repository,[3] where the original 2018 version of the textbook is hosted, indicate that the 2018 edition has been accessed from Germany, France, Norway, Vietnam, Canada, Estonia, the USA, among others.

In addition to the main delivery channels of the OpenUCT repository and the interactive website, the textbook was also disseminated via a number of portals and repositories, such as unglue.it[4], OpenLibra[5], E-Books Directory[6] and Open Textbook Library[7].

Content development and publishing tools

Maria authored the textbook in LaTeX, using the TexShop editor and TexLive. She generated figures using OmniGraffle and references with JabRef. She also used Protégé OWL for making and double-checking exercise files. The website was produced using an HTML editor (first with Taco HTML Edit and then Komodo Edit).

Copyright and licensing

In addition to the text which Maria and her collaborators produced, Maria also adapted third-party CC-licensed content for inclusion in chapter 3 and modified some of her own previously released CC-licensed work for inclusion.

> Maria's decision to openly license her work under a Creative Commons licence helps to facilitate global reach

Maria's *An Introduction to Ontology Engineering* open textbook is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International licence.

Quality assurance and sustainability

In the range of feedback provided by the broader community of scientists that reviewed the content of the textbook, comments included the spotting of typos and taking the effort to inform Maria about it, such as was done by a colleague in Bulgaria. This enabled Maria to correct the new version (and update the list of contributors on the errata page accordingly).

In terms of the sustainability of the textbook, Maria undertook various activities to promote it, including an announcement via a blog post and profiling on social media sites such as LinkedIn.

Status at grant closure

At the close of the grant period, Maria believed that her textbook was a 'proper package deal', stating, 'I hope that some benefits will come from that, whatever they may be.'

Maria stated that she had no plans for another version in the near future. It had taken a lot of work to produce the textbook, which was possible as an activity conducted within a sabbatical period, but which had also substantially affected her research output during that time. She therefore deemed that this level of activity and production would be unfeasible to maintain during a regular academic teaching year.

She also considers that the basics of ontology engineering were relatively stable now and would therefore not require further major updates in the short term.

She highlighted the fact that an interesting next level of work would be the production of a comprehensive textbook (estimated about twice the length) but acknowledged that such a process would require a multi-author approach. This has been discussed in her academic community for the past 10 years, but nothing has come of it to date.

2 http://www.collegepublications.co.uk/computing/?00020
3 https://open.uct.ac.za/handle/11427/28312/statistics
4 https://unglue.it/work/316551/
5 https://openlibra.com/es/book/an-introduction-to-ontology-engineering
6 http://www.e-booksdirectory.com/details.php?ebook=12071
7 https://open.umn.edu/opentextbooks/textbooks/590

Challenges experienced and lessons learned

Maria reflected on both the benefits and the challenges she experienced as part of her open textbook development process.

She realised she was not alone

Maria noted that her participation in the workshops hosted by DOT4D in 2019 provided a comforting realisation that, 'I [was] not the only one spending a disproportionate amount of time on something ideologically good, but what seems to be (perceived to be) a 'time waster''. Indeed, Maria highlighted how there was a good amount of social media feedback on her posts in which she announced the tools. Emails arrived soon after she made the slides available in LaTeX and PowerPoint formats from colleagues around the world. In addition, multiple comments were broadcast 'from a generally critical crowd' on the textbook being freely available.

> The open textbook production process took up a large proportion of her time

Open textbook development takes more time than you think

Maria stated that the open textbook production process took up a large proportion of her time. Maria cautioned others against taking on this endeavour during a regular academic year, as this would clash with the everyday demands and pressures of teaching, research and administration.

Maria stated that, 'what I would do differently next time is to demand buy-out or time off from teaching or administration in order to give the content production more attention. This is unlikely to happen, since it is part of a culture to under-charge teaching workload allocations'.

Don't pin all your hopes on student engagement

It was difficult finding students willing to do the work; and once students were found, they often did not respect deadlines. Due to the reluctance of students to participate in certain tasks, such as aiding parts of the quality assurance process with the typo-finding contest, Maria had to take on these tasks herself in order to meet the specified deadlines.

A series of missed deadlines created a ripple effect within the textbook development process and affected Maria's opportunities to be able to test different aspects of the textbook, such as the OBDA tutorial, due to the fact that there were no students left within the course at the point of its completion to test this tool.

In addition, in 2019, the Ontology Engineering course ran as an MSc course rather than an Honours course (which previously had 30 students), and none of the comparatively few MSc students were interested in taking up any vacation work. Maria suggested that open textbook authors should limit the number of collaborators within the development process, as this led to increased administrative work.

Computer science students aren't very keen to write educational text

While Maria initially assumed the benefits of having students participate in the writing of particular sections of the textbook, the main problem she noted was that computer science students generally disliked the task of writing and that they were inexperienced in the practice of 'writing for learning' due to their lack of teaching experience. Maria highlighted the fact that the precarious engagement with students significantly affected the way in which she had imagined the structure, content and quality of the tutorial for her textbook.

Institutions should fund open textbook development

In her final grant report, Maria expressed her frustrations with the current systems in place, stating that, 'besides aforementioned buy-out to make time for writing the textbook to help making knowledge freely available, I still think money deserves to go also to the open access costs of a journal article reporting on some of the materials'. She stated that if the institution wants academics to publish open textbooks, it should accept that this comes at a price for the content producers and make funding and time available for this task.

> If the institution wants academics to publish open textbooks, it should accept that this comes at a price for the content producers and make funding and time available

Budget

Overview of the original budget submitted to DOT4D as part of 2018 grant application, with actual expenditure.

Budget projected at proposal phase
Software developer: R10,560
OBDA tutorial/exercise: R5,280
OBDA tutorial testing: R230
Multilingual tasks: R9,900
Slides: R3,720
Typo voucher (airtime/data): R300
External expertise: R5,995
Other admin: R930
Bound hardcopies for students: R1,800

DOT4D grant amount: R38,715

Project actual expenditure
Collaborator visit: R5,040
Printing: R1,961
Software development DL plugin + bit multilingual: R6,000
Software development Afrikaans + Spanish ODE: R5,000
Software & content development CLaRO tool + categorising multilingual verbalisation: R4,600
Content development OntoClean tutorial + literature list multilingual ontology verbalisation: R6,600
Content development OBDA with local data: R5,280
Content development slides: R3,720
UCT payroll costs: R199
Typo contest: R300

Total expenditure: R38,700