Open Science Monitoring

Impact Case Study – Polymath

Date: 23/01/2017

Research and Innovation

Author information and acknowledgements

This study was delivered by **RAND Europe**, with the support of **Deloitte** and **Digital Science & Research Solutions**, including **Altmetric.com** and **figshare**. The project was commissioned by the European Commission Directorate-General for Research and Innovation and received funding under specific contract 2015/RTD/A6/SC/PP-02741-2015.[1]

The team at RAND Europe included: Elta Smith (project leader), Salil Gunashekar (project manager), Sarah Parks, Catherine Lichten, Louise Lepetit, Molly Morgan Jones, Catriona Manville and Calum MacLure. The team at Deloitte included: Valentina Cilli, Maruca De Ramón, Vilma Zotou, Benoit Vandresse and Sebastiaan van der Peijl.

Data sources

We would like to thank the following organisations for generously contributing their data for use on the monitor: 101innovations, Altmetric, arXiv, bioRxiv, Clarivate Analytics, F1000Research, figshare, GenBank, Helmholtz-Centre for Environmental Research – UFZ and German Centre for Integrative Biodiversity Research (iDiv), Nature Publishing Group and Palgrave Macmillan, OpenAIRE, PeerJ preprints, Publons, re3data.org, RePEc, SHERPA/RoMEO, SHERPA/JULIET, Taylor & Francis, and Wiley.

Expert panel

We are grateful to our expert panel members who provided helpful advice and guidance: **Aletta Bonn** (German Centre for Integrative Biodiversity Research (iDiv) and Helmholtz Center for Environmental Research (UFZ), Friedrich-Schiller-Universität Jena, Germany); **Thomas Crouzier** (School of Biotechnology, Royal Institute of Technology (KTH), Sweden); **Ivo Grigorov** (National Institute of Aquatic Resources, Technical University of Denmark); **Peter Kraker** (Postdoctoral Fellow, Know-Center Social Computing Group, Graz University of Technology, Austria); **Eleni Malliou** (National Documentation Center, Greece); and **Paul Wouters** (CWTS-Centre for Science and Technology Studies, Leiden University, The Netherlands).

Acknowledgements:

We would like to thank the following individuals for their contributions to the monitor: Susan Guthrie (RAND Europe), Gavin Cochrane (RAND Europe), Alexandra Pollitt (King's College London), and Marion Bywater who acted as our quality assurance reviewers and provided thoughtful and timely comments on our analyses; Jessica Plumridge (RAND Europe) who designed the graphics presented in the online monitor; Monica Hertzman (RAND Corporation) who provided technical assistance in relation to the design of the website; Steven Wooding (RAND Europe) who provided advice and support throughout the study; Vincent Larivière (Observatoire des Sciences et des Technologies) for his insights and advice; Joanna Chataway (University of Sussex) for her advice and support; Rebecca Ioppolo (RAND Europe) who provided research assistant support; Jack Melling (RAND Europe) who helped in the dissemination activities related to the launch of the monitor; and the many individuals and organisations who participated in an online consultation we conducted as part of the development of the monitor.

For more information about this document or the Open Science Monitor, please contact:

Elta Smith
RAND Europe | Westbrook Centre, Milton Road, Cambridge CB4 1YG, United Kingdom
Telephone: +44 (1223) 353 329 | E-mail: eltas@rand.org

[1] Framework Contract 'Provision of Services in the Field of Research Evaluation and Research Policy Analysis', Lot 2: Data collection and performance indicators to monitor European research policy 2010/S 172-262618

Polymath

Summary

The Polymath project (Polymath 2016a) aims to find solutions for complex, unsolved mathematical problems in a collaborative way. Most of the problems are published through a blog post. Each problem has a wiki page that summarises all the knowledge developed for that specific problem, and research and discussion threads, where researchers can post their contributions and discuss solutions.

As of 2016, the Polymath project has published 11 projects (i.e. 11 problems) on combinatorial mathematics.[2] Of these, three have been solved and publications have been produced. One was eventually solved by a researcher who had participated actively in the discussion on Polymath (and acknowledged the relevance of these contributions in arriving at the final solution).

The Polymath project has developed a set of rules on how to publish the problems and conduct the related discussion, and for summarising and consolidating the results (the so-called Polymath general rules). In addition, papers published as part of the Polymath project are authored under the collective pseudonym of 'D.H.J. Polymath' and with a link to the full working record. In case of recognition of specific achievements of one or more contributors, it was decided to use letters of recommendation.

The Polymath project started with the intention of proving whether collaboration in mathematics could lead to faster results. The Polymath1 problem provided an answer to this question, and the collaboration has continued over the years, leading to new problems solved and additional publications. Overall, the focus has remained on combinatorics and the community has not reached the 'massive' dimensions that had been hoped for. Solving the first problem involved the cooperative effort of 27 researchers. More recent problems have involved the contributions of about 40 researchers.

There has been a spin-off in the form of educational activity (the Crowdmath project), in collaboration with the MIT PRIMES program and the Art of Problem Solving. The Crowdmath project is specifically targeted at high school and college students with backgrounds in mathematics and aims to educate young mathematicians to conduct research, particularly through collaborative approaches. It does not focus on new, unsolved problems, instead it focuses on problems for which students can find available materials to support their research.

Background

The Polymath project started in 2009, when University of Cambridge professor and Fields medallist[3] Timothy Gowers set out to understand whether 'massive collaborative

---

[2] Combinatorics is a branch of mathematics which deals with the study of finite or countable discrete structures. It studies the enumeration, combination and permutation of sets of elements and the mathematical relations that characterise their properties. Combinatorics has applications in other areas of mathematics, such as graph theory, coding and cryptography and probability (see Fox 2009).

[3] The Fields Medal is a prize awarded to between two and four mathematicians under 40 years of age at the International Congress of the International Mathematical Union, a meeting that takes place every four years. Together with the Abel award, it is considered one of the most prestigious awards for a mathematician, and it is often described as 'the Nobel prize for mathematics'. Its purpose is to give recognition and support to younger mathematical researchers who have made major contributions to the discipline (see Barran & Weisstein 2017).

mathematics' was a possible path for research in mathematics (Gowers 2009a). He started the Polymath blog (Polymath 2016a), where he could present difficult, unsolved mathematical problems that interested him, and he invited interested readers to share ideas on how to solve those problems (or specific aspects of the problem). Inspiration came from open source software, such as Linux, and initiatives such as Wikipedia, which represented innovative, collaborative ways of developing new products (Gowers & Nielsen 2009).

The Polymath project started with the aim of finding elementary proof of a special case of the density Hales-Jewett theorem (DHJ), which is a central result of combinatorics. This theorem was chosen because finding an elementary proof would require a wealth of new ideas in addition to the (existing) long proof, which relied on cumbersome mathematical calculations. More importantly, it was expected that such an elementary proof would help generate several breakthroughs in the future. The result was achieved more rapidly than expected (in about six weeks) thanks to the input of 27 researchers (ibid.).

Methods and the role of open science

Polymath projects are hosted either on the Polymath project blog or on the blog of a host who is willing to administer and moderate the project. On average, up to 40 researchers participate in each project and make up to 60-70 contributions in total. Use of the blog must follow a set of general rules for the publication of the problems and the comments, and for summarising and consolidating any results (known as the Polymath general rules) (Polymath 2016a).

In general, each Polymath project consists of three main parts, which are set out in the Polymath rules (ibid.):

- **The wiki pages for that project**. These pages, which are hosted on the wiki page for Polymath (2016b), store all the knowledge developed for that project, such as the bibliography, notations, finished arguments, and open problems. Anyone who has registered with the wiki is able to edit these pages.
- **The research thread for the project.** This is known as the 'engine of progress' (ibid.) for the project. Threads are used to provide and analyse ideas and to discuss problem-solving strategies, for example. Longer contributions and computations are published on the wiki and summarised on the research thread.
- **The discussion thread for the project.** This is the place where the project is managed and evaluated and where occasional participants can catch up and offer feedback. At regular intervals, the moderators (or other active participants) summarise the research thread by posting updates to the discussion thread. Occasional participants are encouraged to ask questions or make requests about the status, direction or format of the project.

It is also possible to create other components to a project, depending on the need, such as a reading seminar thread to understand a paper of key importance to the project.

The Polymath project also has a specific set of rules for contributions and comments, namely (Polymath 2016a, 2016b):

- 'Everyone is welcome to participate', though people who have already seen an external solution to the problem are asked to refrain from publishing 'spoilers' during the experiment.
- Comments should be constructive, polite and as easy to understand as possible.
- Comments are not expected to be perfect: it is acceptable for a mathematical thought to be tentative, incomplete, or even incorrect; thus, comments can present only partial solutions or observations.

- 'This is a team effort, not a race between individuals'. Participants are expected to elaborate short observations (or comments to another participant's observation) and to report those online to the entire group, rather than work in isolation. Participants are also encouraged to report partial results or even failures.
- Participants are encouraged to update the wiki, or to summarise progress within threads, for the benefit of others. Linking between the wiki and specific comments in the blogs is encouraged.

Each time a new problem is published, users are free to provide their inputs to the solution of the problem as a whole, but also to specific subproblems or topics. When the discussion on specific aspects of the overall problem develops, threads are added to the main problem so that contributors can focus on the specific area(s) they prefer. The approach of adding threads to the main discussion topic was adopted as a way of supporting the debate, given that one of the main ideas behind the Polymath project was that each individual could contribute with partial answers to the solution of the bigger problem. In addition, it makes it easier for contributors to post their input under the relevant discussion point.

The process is structured as an open, collaborative discussion that has led participants to freely contribute in any way they feel capable, from large demonstrations to minor comments. There is no censorship of the scientific value of single contributions.[4] All comments are considered constructive, even when the argument proposed proves to be wrong or not relevant to the problem tackled. Scientific progress is considered a non-linear process, of which error is an inherent part; therefore 'wrong' contributions are still important as they can help identify misconstructions and thus elaborate correct solutions and/or identify new problems and original answers.

The Polymath project makes the research process entirely open and visible, including the way research ideas change over time. It demonstrates how progress can be achieved through the aggregation and progressive refinement of many partial contributions (essentially making the entire 'trial and error' process fully transparent). The initiative thus provides an example of how the research process and results can be made accessible to other researchers and citizens in general. Moreover, participants are free to suggest mathematical problems and coordinate them as collaborative projects via the Polymath project wiki page.

Access to the Polymath project is not restricted to academics, and access to the blogs and the wiki page does not require academic credentials. The blogs (and the wiki page) have moderators checking messages before their publication so as to avoid publication of off-topic or irrelevant contributions. However, the moderation work is minimal, and according to the founders and main contributors, spam and off-topic discussions are negligible (Gowers & Nielsen 2009). This makes it a good example of how scientific progress can be made in an open and collaborative manner.

Of the 11 problems published to date, problems 1, 4 and 8 were solved via the collaborative approach supported by the platform and have generated scientific publications.

Polymath1 was the first problem published, and it was the one which initiated the platform. As the project took form, two main threads of discourse emerged. The first thread, which was run in the comments of Gowers' blog, continued with the original goal of finding a combinatorial proof. The second thread, which was run in the comments of Terence Tao's blog, focused on calculating bounds on density of Hales-Jewett numbers and Moser numbers for low dimensions. After six weeks, Gowers announced on his blog that the problem was

[4] A basic assessment scheme has been introduced. Users can signal the relevance of the different contributions by using the 'thumbs up' and 'thumbs down' signs.

‘probably solved’ (Nielsen 2009a), even though work continued on both Gowers' and Tao's threads well into May 2009, some three months after the initial announcement. Both threads of the Polymath1 project have been successful, producing two papers published under the pseudonym D.H.J. Polymath.

Polymath4 was started to identify a deterministic way to find prime numbers, in July 2007. It generated several discussion threads (up to 7) in different periods of time, two of which (threads V and II) are still active. A solution to the main problem was found in 2010, and a research paper was published.

The Polymath8 project was proposed to improve the bounds for small gaps between primes. It has two components, Polymath8a on bounded gaps between primes, and Polymath8b, focusing on bounded intervals with many primes. Both of them have been successful, producing two new papers published under the pseudonym D.H.J. Polymath.

For some of the other projects, while no official solution was declared, the online discussion led to what researchers have defined as ‘fruitful’ discussion (Gowers 2015), namely, interesting points from the online debate that were eventually used by individual researchers for their own work. For instance, even though problem 5 was not solved, in September 2015, Terence Tao, one of the participants of Polymath5, solved the problem in a pair of papers. One of the two papers showed how the result achieved by the author, combined with some of the arguments elaborated by Polymath5, would provide a complete solution to the problem.

In addition, in March 2016, the Polymath project began sponsoring a ‘crowdmath’ project in collaboration with the Massachusetts Institute of Technology (MIT) PRIMES program (Polymath 2016a; MIT Department of Mathmatics 2016) and Art of Problem Solving (AoPS) (2016). The Crowdmath project targets high school and college students with backgrounds in mathematics and aims to educate young mathematicians in how to conduct research and, in particular, to do so collaboratively, thus supporting the growth of ‘young (collaborative) researchers’. The idea behind this project is the same as for the Polymath project – that is, that massive collaboration in mathematics is possible and fruitful for elaborating solutions to mathematical problems. The Crowdmath project is carried out in cooperation with the MIT PRIMES programme,[5] a free, year-long after-school research programme for high school students, organised by MIT and hosted on the AoPS platform, which provides interactive educational opportunities for mathematics students (AoPS 2016).

The Crowdmath project targets high school and college students from all over the world with a background in advanced mathematics. Middle school students can join in, but only if they have an extremely advanced mathematical background relative to their years of schooling. PhD students and professors are encouraged to participate as advisors and mentors and to provide support and guidance to young students.[6]

The Crowdmath project works in much the same way as the Polymath project. For each problem, there are both a wiki page and threads where users can post their comments. In

[5] The PRIMES programme, which stands for Program for Research in Mathematics, Engineering and Science for High School Students, focuses on research in the fields of mathematics and engineering, and targets high school students with a strong background and interest in those topics (participation from middle school students is also possible). Students apply to the MIT PRIMES programme by submitting research projects, for which they receive not only tutoring but also financial support. Overall, all the projects supported by the MIT PRIMES programme have generated at least one academic publication and have enabled students to receive college scholarships (see MIT Mathematics Department 2016).

[6] There were a total of nine PhD students and professors (mostly from MIT) engaged as mentors and advisors to the project as of October 2016.

addition, mentors provide reference documents to help students find a solution. As in the case of the Polymath project, any academic paper resulting from the Crowdmath research will be published under a collective pseudonym.

The project advisors select problems suited to all participating students. Therefore, the selection favours problems that would be acceptable for high school and college students with a background in advanced mathematics, rather than new or unsolved problems (Geneson, pers. comm. 2016). The availability of materials on the problems is a crucial element for the selection. The problems published so far have focused on discrete mathematics and have been analysed within the MIT PRIMES project, so that students can have supporting materials.

The first Crowdmath project began in March 2016, and so far seven problems have been published. Mentors review individual contributions. Students are assessed based on the relevance of the contributions provided. The initial target for participation was about 20 students, according to the researchers involved. After about six months, 33 students were participating, and mentors and advisors interviewed were confident that the number would increase further in the following months (but had set no specific target) (ibid.).

Outputs and findings

While up to 40 researchers now participate in solving each of the Polymath problems, making up to 60-70 contributions, this falls short of the ambition of 'massive' collaboration in mathematical research. The issue of the scalability of the Polymath approach has been debated by the creators of the project, but no specific solutions have been identified (Nielsen 2009b).

In parallel to the initial mathematical problem (Polymath1), Tim Gowers published a series of posts on the nature and possible mechanisms of 'massive collaborative mathematics' (Gowers 2009a). This led to a parallel debate in the combinatorics research community on the relevance of the Polymath project for establishing a paradigm for collaborative research in mathematics.[7]

The debate focused on the characteristics that the Polymath projects should have (e.g. degree of complexity, discussions around solutions to subproblems) (Gowers & Nielsen 2009) and on the incentives for researchers to work collaboratively (e.g. on recognition of the contributions to the final solution of the problem, authorships of publications) (Gowers 2009b). Some of the rules developed for publishing a problem and contributing to its solution (such as the wiki page, the research and discussion threads, mechanisms to include 'late' contributors) are a direct result of such discussion and of the experiences and feedback from the first users of the blog.

As the collaborative process made it impossible to account for the specific relevance of each contribution, a solution for the authorship of scientific publications had to be found (Nielsen 2012). This led to the decision to sign academic papers published reporting the results of successful Polymath projects under the collective pseudonym D.H.J. Polymath, with a link to the full working record (Polymath 2016a). Where it was appropriate to recognise the specific achievements of one or more contributors, it was decided to use letters of recommendation, as is done in particles physics (where papers can have hundreds of authors) (Gowers & Nielsen 2009). This decision does not seem to have prevented researchers from providing

[7] The original post on Glower's blog generated 212 comments (including author's replies) on if and how such an experiment could work. It also generated a series of posts on the personal blogs of other main contributors of the Polymath project (such as Profs. Tao and Nielsen) and several publications on the functioning of the project.

inputs to the Polymath projects, as the number of individual participants has increased slightly over the years.[8] In addition, indirect contributions provided by Polymath to solutions published by individual researchers (or groups of researchers) are acknowledged.

The work on Polymath5 is one such example (Gowers 2015). The work carried out under the Polymath5 project helped to shed the light on some of the aspects of what is known as the Erdős discrepancy problem. Terence Tao used this work a few years later, in combination with recent developments in number theory. According to Tim Gowers, Polymath5 created 'a sort of penumbra around the problem, which had the effect that when these remarkable results in number theory became available, you 'just' have to find a function with certain properties and you are done' (ibid.).

In his solution of the Erdős discrepancy problem, Terence Tao uses two arguments (i.e. the Fourier-reduction and the Elliot conjecture) that had already been discussed in the polymath project, and he provided a direct reference to that discussion. Therefore, again in the words of Tim Gowers, 'it seems likely that in some sense polymath5 played a role in provoking Terry to have the thoughts he did' (ibid.).

The collaborative approach applied in the field of combinatorics mathematics has been praised by the research community in combinatorics and in mathematics in general (Gowers & Nielsen 2009). Some of the problems have been resolved and published, while others have provided relevant inputs for subsequent solutions (such as the work on the Polymath5 problem) (Rehmeyer 2009).

However, doubts have been expressed about the scalability of this approach to other branches of mathematics or to scientific research generally. On the one hand, one of the likely success factors of the Polymath projects has been its focus on combinatorics, which has a relatively small, quite connected community (Ellenberg 2009). In addition, research in combinatorics is not accompanied by large economic rewards. Consequently, the incentives for direct academic credit or patents deriving from research are quite low, and this removes a possible obstacle to cooperation (Gowers & Nielsen 2009). Finally, while an academic background is not required to join the discussion on the Polymath project, most contributors are from academia (researchers or PhD students).

Overall, five academic articles (Polymath, D.H.J. 2010, 2012, 2014a, 2014b, 2014c) have been published under the Polymath collective pseudonym as a result of the solutions found to Polymath problems 1, 4 and 8. Journal editors for the sixth (Tao et al. 2012) required the authors to publish under their real names although the arXiv version uses the Polymath pseudonym.

Impacts

The Polymath project started with the idea of proving whether collaboration in mathematics could lead to faster results. The Polymath1 problem proved this to be the case, and the collaboration has continued over the years, leading to new problems being solved and additional publications being published. The impact has been confined to combinatorics, and the impact on the wider mathematics community has remained limited, so that the hope of achieving 'massive' dimensions has not been realised. There has, however, been an impact on the future research community through the sponsorship of Crowdmath.

[8] From 27 contributing researchers for the first problem to about 40 researchers for more recent problems.

Sources

Interviews

We gratefully acknowledge the support of the following interviewee, who provided information for this case study: Jesse Geneson, Crowdmath project.

Bibliography

Art of Problem Solving (AoPS). 2016. 'CrowdMath.' As of 18 January 2017: http://www.artofproblemsolving.com/polymath/mitprimes2016

Barran, M., & E.W. Weisstein. 2017. 'Fields Medal.' MathWorld – A Wolfram Web Resource. As of 18 January 2017: http://mathworld.wolfram.com/FieldsMedal.html

Ellenberg, J. 2009. 'The Ninth Annual Year in Ideas: Massively Collaborative Mathematics.' *New York Times*.

Fox, J. 2009. Lecture 1: Introduction, MAT 307: Combinatorics. As of 18 January 2017: http://math.mit.edu/~fox/MAT307-lecture01.pdf

Gowers, T. 2009a. 'Is Massively Collaborative Mathematics Possible?' 27 January. As of 18 January 2017: https://gowers.wordpress.com/2009/01/27/is-massively-collaborative-mathematics-possible/

———. 2009b. 'Polymath1 and Open Collaborative Mathematics.' 10 March. As of 18 January 2017: https://gowers.wordpress.com/2009/03/10/polymath1-and-open-collaborative-mathematics/

———. 2015. 'EDP28 – Problem Solved by Terence Tao!' 20 September. As of 18 January 2017: https://gowers.wordpress.com/2015/09/20/edp28-problem-solved-by-terence-tao/

Gowers, T., & M. Nielsen. 2009. 'Massively Collaborative Mathematics.' *Nature* 461: 879–81.

Harriss, E. 2009. 'Polymath.' March 14. As of 18 January 2017: https://maxwelldemon.com/2009/03/14/polymath/

IAS (Institute for Advanced Study). 2010. 'Mathematical Advances: Lone or Massively Collaborative Endeavors?' from IAS Institute Letter for fall 2010 based on a discussion organized by IAS fall 2010. As of 18 January 2017: https://polymathprojects.files.wordpress.com/2011/03/polymathias.jpg

MIT Department of Mathematics (Massachusetts Institute of Technology). 2016. 'PRIMES: Program for Research in Mathematics, Engineering and Science for High School Students.' As of 18 January 2017: http://math.mit.edu/research/highschool/primes/index.php

Nielsen, M. 2009a. 'The Polymath Project: Scope of Participation.' 20 March. As of 18 January 2017: http://michaelnielsen.org/blog/the-polymath-project-scope-of-participation/

———. 2009b. 'On Scaling Up the Polymath Project.' 25 March. As of 18 January 2017: http://michaelnielsen.org/blog/on-scaling-up-the-polymath-project/

———. 2012. *Reinventing Discovery: The New era of Networked Science.* Princeton, NJ: Princeton University Press.

Polymath. 2016a. 'The Polymath Blog.' As of 18 January 2017: https://polymathprojects.org/

———. 2016b. 'Main Page.' As of 18 January 2017: http://michaelnielsen.org/polymath1/index.php?title=Main_Page

Polymath, D.H.J. 2010. 'Density Hales-Jewett and Moser Numbers, An irregular Mind.' *Bolyai Society Mathematical Studies* 21: 689–753 (from the Polymath1 project).

———. 2012. 'A New Proof of the Density Hales-Jewett Theorem.' *Annals of Mathematics, Second Series,* 175 (3): 1283–1327 (from the Polymath1 project).

———. 2014a. 'New equidistribution Estimates of Zhang Type.' *Algebra & Number Theory* 9 (8): 2067–2199 (from the Polymath8 project).

———. 2014b. 'Variants of the Selberg Sieve, and bounded Intervals Containing Many Primes.' *Research in the Mathematical Sciences* 1 (12) (from the Polymath8 project).

———. 2014c. 'The "Bounded Gaps Between Primes" Polymath Project: A Retrospective Analysis.' *Newsletter of the European Mathematical Society* 94: 13–23.

Rehmeyer, J. 2009. 'Mathematics by Collaboration.' *ScienceNews*, 8 December.

Tao, T., E. Croot III & H. Helfgott. 2012. 'Deterministic Methods to Find Primes.' *Mathematics of Computation* 81 (278): 1233–46 (from the Polymath4 project).