Curating Online Mathematical Resources

Pieter Belmans

There have been contributions in the Early Career Section discussing your online presence as a mathematician [Kri19], and how to use social media [Ric20]. And we've been told we need to tell our own stories as mathematicians, and make mathematics *outward facing*, by writing about mathematics for general audiences [Ell19].

In this article I want to discuss another form of science communication, that of making mathematical results interactively accessible on the web. Two ways of doing this I personally know well are:

1. making large-scale mathematical texts available in an online-first way,
2. turning mathematical classification results into interactive websites.

To illustrate the first point I will give a brief history of the Stacks project. This is nowadays a household name in algebraic geometry. But also if you are not already familiar with it, it provides an interesting case study appealing to all mathematicians. To illustrate the second point I will showcase some examples of how such interactive websites can be used as both a *communication device* and a *research tool*, with a plea to not forget the communication aspect.

I hope this article acts as an invitation to *discover* and *use* online mathematical resources. There are likely websites in your research area that you don't know about yet, or which have ways of being used that you haven't thought about yet! But it is also an invitation to the more tech-savvy mathematicians to *create* new online mathematical resources. One can find great pleasure (at least, I do) in this particular form of science communication. By showcasing some examples I hope to pique your interest.

The Stacks Project

This is an online reference work in algebraic geometry, whose history I would like to briefly discuss, as it can teach us hopefully something about how (not) to approach collaborative mathematics on the internet.[2]

For those outside algebraic geometry circles, the Stacks project describes itself as:

> an open source textbook and reference work on algebraic geometry

As a *"book"* it is a hefty tome of more than 7200 pages at the time of writing, and it regularly is extended. It deals with the foundations of algebraic geometry and the (rather technical) theory of algebraic stacks, a short introduction to which can be read in the *Notices* "What is...?" series [Edi03]. As a *website* it can be accessed at `https://stacks.math.columbia.edu` and this is the recommended way of reading it. But how did we get to this point?

The Germ of an Idea

Inspired by open source programming projects, and the Linux kernel in particular, Aise Johan de Jong sent an email to a group of colleagues on **June 20, 2005**, proposing a mathematical collaboration. He wanted to collectively write a book that could serve as an introduction to the theory of algebraic stacks. But his proposal came with an important caveat, making it different from the usual process of writing such a text:

- established researchers would make an effort on an initial manuscript,
- this would then be published online,
- everyone (colleagues, but also students) would submit corrections and new contributions

and all of this would be coordinated through an email list.

It would also be essential that everything is published using an *open source license.* This is common for programming projects, making the source code public and inviting everyone to contribute. Yet for mathematical texts this was unheard of. Wikipedia, which started in 2001 and became the most popular online reference in 2005, was the closest parallel. But Wikipedia excels in its breadth, not its depth, and only gives information already available elsewhere; Johan's suggestion would focus on a single topic within a single area of mathematics, and discuss a topic for which not many textbooks existed at the time.

People responded with enthusiasm, but there was also a pessimistic voice [Buza]:

> "Hi Johan et al. Here's why I think this project won't work *yet*, and feel free to prove me wrong!"—*Kevin Buzzard, 2005*

We can summarise his arguments from the email as follows:

- the Linux kernel wasn't started by a team of programmers; Linus Torvalds single-handedly wrote the first version of the kernel, and only when that turned out to be viable did other people get involved;
- there are as many opinions about content, the approach to take or notation to use, as there will be mathematicians involved, leading to conflicting opinions and lots of communication overhead.

He also suggested a solution:

> "I also conjecture that essentially the *only* way that my pessimistic prediction above will fail is that Johan himself stops doing his

Pieter Belmans is an assistant professor of mathematics at the University of Luxembourg. His email address is `pieter.belmans@uni.lu`.

[2]*A longer version of this (in Dutch) was published earlier by de Jong and the author [BdJ20], parts of which are reproduced here with permission.*

DOI: https://dx.doi.org/10.1090/noti2344

> research and spends 6 months writing version 0 of the *entire book*. I don't know whether he is prepared to do this. But I feel that this is absolutely key because it is the unique way to give concrete direction to the project.
>
> [...] but if he just sits down and writes several hundred pages by himself then whenever anyone asks a question like this, he can just point to the book and say 'that's what I think it should be, you go ahead and add some more stuff if you think differently'."—*Kevin Buzzard, 2005*

This proposed solution actually nicely parallels Kevin's role in formalisation of mathematics: starting in 2017 he has become an important figure in this large-scale open source project. He writes [Buzb]

> "So I have decided to stop attempting to generate new mathematics, and concentrate instead on carefully checking 'known' mathematics on a computer."—*Kevin Buzzard, 2020*

But that is the topic for a different article.

A False Start, and Then the Real Start

After a lengthy discussion of the idea in July 2005, no writing happened. A preliminary structure was built by Johan in September and October of that year, with mostly placeholder sections to make the intended structure visible. The next few weeks saw some "patches" by contributors of the mailing list, mostly PhD students of Johan. By early 2006 a 70 page document existed, and that's where it stopped.

The *real* starting date for the Stacks project is **May 21, 2008**. That is the date when Johan did the "Initial commit" in Git,[3] starting with the text from 2006. Some important decisions were made:

- Johan would do most of the writing and editorial work,
- the goal would be less on writing an introductory textbook, and more on writing an exhaustive reference work.

This latter point follows the tradition of the famous reference works *Éléments de Géométrie Algébrique* (EGA), written in the 1960s by Grothendieck with the help of Dieudonné, developing the theory of *schemes*. These span four books (with 13 planned in total) published in eight volumes, but the theory of algebraic stacks did not exist at that time. The Stacks project was to become an EGA of the 21st century, encompassing both the theory of schemes and stacks.

Robust Referencing: The Tags System

For a slowly evolving mathematical reference work it is important to have a robust referencing system. You don't want Theorem 21.4.2 to become Theorem 21.5.2 when you've added a new Subsection 21.4.

The solution, suggested by Cathy O'Neil, was the *tags system*. A tag is a unique and stable identification of a result in the Stacks project, consisting of four numbers and letters. For example, the definition of an algebraic stack can be found as tag 026N. Tags are regularly assigned to new statements in the Stacks project, and follow a roughly chronological order, instead of the order in which they appear in the text.

This system was introduced on **May 15, 2009**, when the Stacks project already contained 3026 tags. At the time of writing there are 20536 tags. So there's plenty of room to go, with 10 digits and 25 letters (as it is too easy to confuse the numeral 0 with the letter O) giving $35^4 = 1500625$ possible tags.

But it is of course still possible (and for many readers desirable) to read the Stacks project using the more familiar a.b.c numbering format, as this gives an important sense of orientation in the text which the tags system cannot provide.

Going Online

In the beginning the only way to read the Stacks project was through a collection of pdfs. After the tags system was introduced a rudimentary lookup system was in place, where through a basic form you could look up the TeX code for a tag. To make this experience better, Johan proposed a "summer project" on **May 6, 2012**, to improve the lookup system and make commenting possible. The only volunteer was the author of this article, and after some procrastination during his exams, the first version of the website appeared on **July 19, 2012**, with a "feature complete" version going live on **July 30, 2013**.

In 2017 it was decided to overhaul the entire infrastructure running the website. It was very ad hoc and only worked for the Stacks project, but it was clear that other projects could benefit from a similar infrastructure. This led to the creation of *Gerby*, a somewhat flexible system to run the Stacks project and other websites using a tags system. Together with an overhaul of the layout, making the website accessible on mobile devices, the second version came live on **May 21, 2018**. Exactly 10 years after the birth of the Stacks project.

The Deligne–Mumford Promise

Another interesting date in the history of the Stacks project is **August 14, 2017**. This is the day that tag 0E9C was added. It corresponds to the following theorem (which you are not necessarily supposed to understand).

> **Theorem 1** (Deligne–Mumford). *Let $g \geq 2$. The algebraic stack $\overline{\mathcal{M}}_g$ is a Deligne–Mumford stack, proper and smooth over Spec $\mathbb{Z}$. Moreover, the locus $\mathcal{M}_g$ parametrizing smooth curves is a dense open substack.*

This theorem of Deligne and Mumford was proven in 1969, and appears as Theorem 5.2 in [DM69]. This is the

[3] *See* `https://github.com/stacks/stacks-project/commit/3d32323ff9f1166afb3ee0ecaa10093dc764a50d`.

article that introduced algebraic stacks. But here is an important quote from that article:

> "Full details on the basic properties and theorems for algebraic stacks will be given elsewhere."—*Pierre Deligne and David Mumford, in* [DM69], 1969

The foundations of algebraic stacks did not exist at that time, and whilst no one doubted the correctness of the many technical details that go into the development of the theory (and in the meantime many of those details appeared in various other works), the first time the full theory of algebraic stacks was written down was in the Stacks project, almost 60 years later.

The proof of the Deligne–Mumford theorem in the Stacks project takes 6610 tags, almost a third of all tags, making it the most complex theorem in the Stacks project, requiring lots of foundational work.

Other Examples, and a Conclusion

As mentioned before, the current infrastructure running the Stacks project can be used for other similar projects. The main example is Jacob Lurie's Kerodon, whose tagline is:

> an online resource for homotopy-coherent mathematics

Jacob has several thousands of pages of mathematics available as pdfs on his website, and he is currently reworking things to make them much more accessible using a tags-based system. The current version can be read at `https://kerodon.net` and it stands at more than 3000 tags, on more than 1000 pages.

These are of course far from the only large-scale mathematical texts out there. It is exciting to see that for permanently evolving works an online-first format really is the best choice, and that web technology is in a sufficiently mature state to make this feasible!

Online Interactive Classifications

Chemistry has the periodic table. Biology has the taxonomy of species in domains, kingdoms, phyla, etc. Geography has atlases. These are all well-known classifications and descriptions of the real world, in a very visual way. At least for me these were responsible for a significant part of the appeal of these sciences.

In mathematics, starting from just axioms, sometimes a classifiable type of objects emerges. One of the most famous instances is the *classification of finite simple groups*, a decades-spanning work that shows that every finite simple group is isomorphic to either a group from one of three infinite lists (cyclic groups of prime order, alternating groups of degree at least 5, or groups of Lie type) or one of 26 sporadic groups. This classification led to the *ATLAS of Finite Groups*, a large (420mm by 300mm, or 1.4 square foot) cherry red tome, giving information about finite simple groups, first published in 1985. In the year 2021 this information can be accessed much more conveniently via `http://brauer.maths.qmul.ac.uk/Atlas/v3/`.

Large-scale Mathematical Websites

There exist other large-scale efforts to make lots of mathematical data readily accessible on the internet. Two important instances that are discussed in this Early Career section are:

- The L-Functions and Modular Forms Database, `https://www.lmfdb.org`,
- The Online Encyclopedia of Integer Sequences, `https://oeis.org`.

A very incomplete and biased continuation of this list would be:

- GroupNames, `http://groupnames.org`,
- The Complexity Zoo, `https://complexityzoo.net`,
- Digital Library of Mathematical Functions, `https://dlmf.nist.gov`.

These are efforts by many people, and they are more research tools than communication devices. If a layperson sees them for the first time, there is not much for them to latch on to. For that reason I would like to switch gears a bit and highlight a different kind of online mathematical resource.

Outward-facing Mathematical Websites

It is in some cases possible to make a research tool that is *also* a communication device. We all know the trope of a mathematician's website looking like it was made in 1994. Or maybe you are unaware of this, and are such a mathematician yourself. But in 2021 the availability of excellent open-source tools for web development makes it possible for more and more mathematicians (with the appropriate background) to experiment with making their area of mathematics visible on the internet.

So now we come to the part of this article in which I make three suggestions:

- If you maintain a mathematical database, make it useable by nonexperts, and make it aesthetically pleasing to make a good first (and hopefully second, and so on) impression on anyone seeing it.
- If you do not (yet) maintain a mathematical database, consider using an interactive website as both a research tool and a communication device. There is likely a classification result in your area of mathematics which would make for a good website![4]
- If the first two suggestions do not apply to you, *look up* and *use* existing online mathematical resources. There are already more out there than you'd imagine!

[4] *As an added benefit, you'll learn a lot about the topic and you'll come up with new research questions. You'll understand what is already known, what should be known but isn't explicitly available, and what is wide open. It's like writing a survey article, but more fun!*

Let me list my own attempts at outward-facing mathematical websites, which hopefully supports the point I'm trying to make about how one could approach online mathematical resources:

- Grassmannian.info, `https://grassmannian.info`, on the geometry of generalised Grassmannians,
- le superficie algebriche, `https://superficie.info`, on the classification of smooth projective surfaces (joint with Johan Commelin),
- Fanography, `https://fanography.info`, on the classification of Fano 3-folds.

These all try to make information which is scattered over many books and references accessible in a convenient way, in a (hopefully) aesthetically pleasing way, that (hopefully) speaks to nonexperts. Of course, the layperson is not necessarily able to fully interact with these, lacking the mathematical background. Going back to chemistry's periodic table: if it weren't part of the high school curriculum, it would also need the guidance of an expert to make this visually appealing classification understandable.[5]

So repeating Ellenberg's appeal from [Ell19]: we need more outward-facing math, and initiatives like online interactive classifications are a great tool for this as they use a format already familiar from other sciences. And then I haven't even talked about the tools currently available to make mathematical visualisations!

References

[BdJ20] P. Belmans and A. J. de Jong, *Het Stacks project*, Nieuw Archief voor Wiskunde **21** (2020), no. 5, 167–173.

[Buza] K. Buzzard, email from Kevin Buzzard to the algebraic geometry mailing list on July 22, 2005, `http://math.columbia.edu/~dejong/wordpress/wp-content/uploads/2010/10/Kevin-Buzzard`.

[Buzb] K. Buzzard, *The future of mathematics?*, Presentation at Formal Methods in Mathematics, Pittsburgh, `https://www.andrew.cmu.edu/user/avigad/meetings/fomm2020/slides/fomm_buzzard.pdf`.

[DM69] P. Deligne and D. Mumford, *The irreducibility of the space of curves of given genus*, Inst. Hautes Études Sci. Publ. Math. **36** (1969), 75–109. MR262240

[Edi03] D. Edidin, *What is...a stack?*, Notices Amer. Math. Soc. **50** (2003), no. 4, 458–459.

[Ell19] J. Ellenberg, *Outward-facing mathematics: A pitch*, Notices Amer. Math. Soc. **66** (2019), no. 3, 354–356.

[Kri19] H. Krieger, *Cultivating an online presence for the academic job market*, Notices Amer. Math. Soc. **66** (2019), no. 6, 853–854.

[Ric20] D. Richeson, *Communicating mathematics using social media*, Notices Amer. Math. Soc. **67** (2020), no. 2, 182–184.

[5]*For the classification of finite simple groups Ivan Andrus has used the periodic table format to create a periodic table of finite simple groups; see* `https://irandrus.files.wordpress.com/2012/06/periodic-table-of-groups.pdf`. *It would be great to make this interactive, as a gateway to the online ATLAS.*

Pieter Belmans

Credits

Photo of Pieter Belmans is courtesy of Ann Borrey.