Research Debt

AUTHORS

Chris Olah
Shan Carter

AFFILIATIONS

Google Brain
Google Brain

PUBLISHED

March 22, 2017

DOI

10.23915/distill.00005

Achieving a research-level understanding of most topics is like climbing a mountain. Aspiring researchers must struggle to understand vast bodies of work that came before them, to learn techniques, and to gain intuition. Upon reaching the top, the new researcher begins doing novel work, throwing new stones onto the top of the mountain and making it a little taller for whoever comes next.

Mathematics is a striking example of this. For centuries, countless minds have climbed the mountain range of mathematics and laid new boulders at the top. Over time, different peaks formed, built on top of particularly beautiful results. Now the peaks of mathematics are so numerous and steep that no person can climb them all. Even with a lifetime of dedicated effort, a mathematician may only enjoy some of their vistas.

People expect the climb to be hard. It reflects the tremendous progress and cumulative effort that's gone into mathematics. The climb is seen as an intellectual pilgrimage, the labor a rite of passage. But the climb could be massively easier. It's entirely possible to build paths and staircases into these mountains. [1] The climb isn't something to be proud of.

The climb isn't progress: the climb is a mountain of debt.

The Debt

Programmers talk about technical debt: there are ways to write software that are faster in the short run but problematic in the long run. Managers talk about institutional debt: institutions can grow quickly at the cost of bad practices creeping in. Both are easy to accumulate but hard to get rid of.

Research can also have debt. It comes in several forms:

- **Poor Exposition** – Often, there is no good explanation of important ideas and one has to struggle to understand them. This problem is so pervasive that we take it for granted and don't appreciate how much better things could be.

- **Undigested Ideas** – Most ideas start off rough and hard to understand. They become radically easier as we polish them, developing the right analogies, language, and ways of thinking.

- **Bad abstractions and notation** – Abstractions and notation are the user interface of research, shaping how we think and communicate. Unfortunately, we often get stuck with the first formalisms to develop even when they're bad. For example, an object with extra electrons is negative, and [1].

- **Noise** – Being a researcher is like standing in the middle of a construction site. Countless papers scream for your attention and there's no easy way to filter or summarize them. [2] We think noise is the main way experts experience research debt.

The insidious thing about research debt is that it's normal. Everyone takes it for granted, and doesn't realize that things could be different. For example, it's normal to give very mediocre explanations of research, and people perceive that to be the ceiling of explanation quality. On the rare occasions that truly excellent explanations come along, people see them as one-off miracles rather than a sign that we could systematically be doing better.

Interpretive Labor

There's a tradeoff between the energy put into explaining an idea, and the energy needed to understand it. On one extreme, the explainer can painstakingly craft a beautiful explanation, leading their audience to understanding without even realizing it could have been difficult. On the other extreme, the explainer can do the absolute minimum and abandon their audience to struggle. This energy is called interpretive labor [2, 3].

Many explanations are not one-to-one. People give lectures, write books, or communicate online. In these one-to-many cases, each member of the audience pays the cost of understanding, even though the cost of explaining stays the same. [3] As a result, the cost of understanding has a multiplier in the interpretive labor tradeoff — sometimes a huge multiplier. [4]
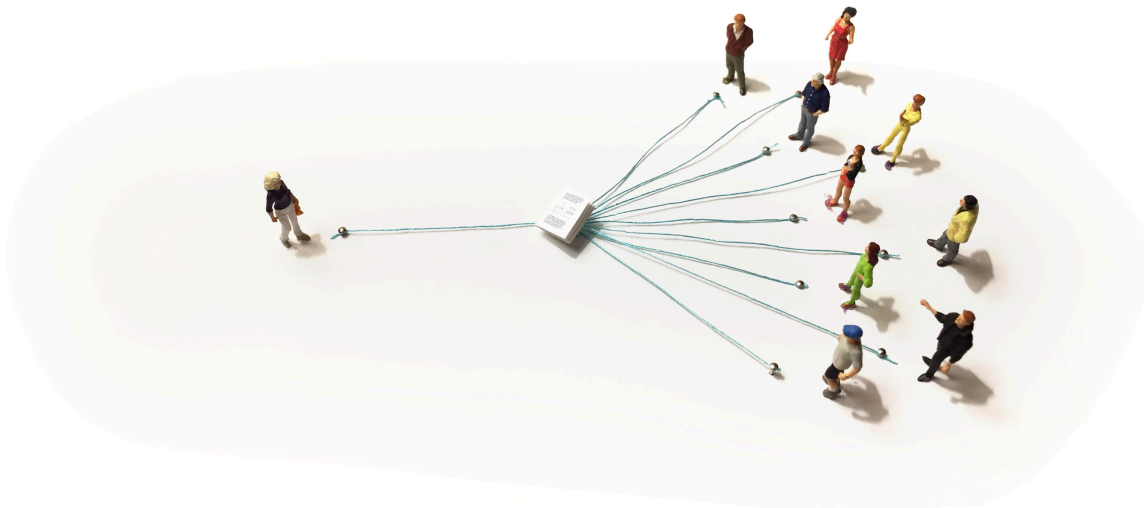
In research, we often have a group of researchers all trying to understand each other. Just like before, the cost of explaining stays constant as the group grows, but the cost of understanding increases with each new member. At some size, the effort to understand everyone else becomes too much. As a defense mechanism, people specialize, focusing on a narrower area of interest. The maintainable size of the field is controlled by how its members trade off the energy between communicating and understanding.

Research debt is the accumulation of missing interpretive labor. It's extremely natural for young ideas to go through a stage of debt, like early prototypes in engineering. The problem is that we often stop at that point. Young ideas aren't ending points for us to put in a paper and abandon. When we let things stop there the debt piles up. It becomes harder to understand and build on each other's work and the field fragments.

Clear Thinking

It's worth being clear that research debt isn't just about ideas not being explained well. It's a lack of digesting ideas – or, at least, a lack of the public version of ideas being digested. [5] It's a communal messiness of thought.

Developing good abstractions, notations, visualizations, and so forth, is improving the user interfaces for ideas. This helps both with understanding ideas for the first time and with thinking clearly about them. Conversely, if we can't explain an idea well, that's often a sign that we don't understand it as well as we could.

It shouldn't be that surprising that these two largely go hand in hand. Part of thinking is having a conversation with ourselves.

Research Distillation

Research distillation is the opposite of research debt. It can be incredibly satisfying, combining deep scientific understanding, empathy, and design to do justice to our research and lay bare beautiful insights.

Distillation is also hard. It's tempting to think of explaining an idea as just putting a layer of polish on it, but good explanations often involve transforming the idea. This kind of refinement of an idea can take just as much effort and deep understanding as the initial discovery.

This leaves us with no easy way out. We can't solve research debt by having one person write a textbook: their energy is spread too thin to polish every idea from scratch. We can't outsource distillation to less skilled non-experts: refining and explaining ideas requires creativity and deep understanding, just as much as novel research.

Research distillation doesn't have to be you, but it does have to be us.

Where are the Distillers?

Like the theoretician, the experimentalist or the research engineer, the research distiller is an integral role for a healthy research community. Right now, almost no one is filling it.

Why do researchers not work on distillation? One possibility is perverse incentives, like wanting your work to look difficult. Those certainly exist, but we don't think they're the main factor. [6] Another possibility is that they don't enjoy research distillation. Again, we don't think that's what's going on.

Lots of people want to work on research distillation. Unfortunately, it's very difficult to do so, because we don't support them. [7]

An aspiring research distiller lacks many things that are easy to take for granted: a career path, places to learn, examples and role models. Underlying this is a deeper issue: their work isn't seen as a real research contribution. We need to fix this.

An Ecosystem for Distillation

If you are excited to distill ideas, seek clarity, and build beautiful explanations, we are letting you down. You have something precious to contribute and we aren't supporting you the way we should.

The Distill Ecosystem is an attempt to better support this kind of work. Right now, it has three parts:

- The Distill Journal – A venue to give traditional validation to non-traditional contributions.

- The Distill Prize – A $10,000 prize to acknowledge outstanding explanations of machine learning.

- Distill Infrastructure – Tools for making beautiful interactive essays.

This is just a start: there's a lot more that needs to be done. A complete ecosystem for this kind of work needs several other components, including places where one can learn these skills and reliable sources of employment doing this kind of work. We're optimistic that will come with time.

Further Reading
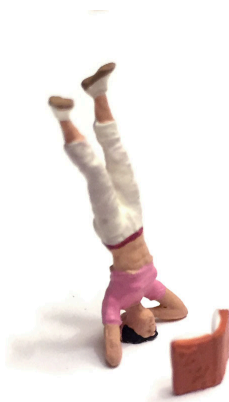
- **Visual Mathematics:** Several mathematicians have made remarkable efforts to visually explain certain topics. Needham's tour-de-force [5] is particularly striking, but there are several lovely examples of new clarity being brought to a topic by visually reformulating it [6, 7, 8] and, on a smaller scale, a plethora of visual proofs.

- **Explorable Explanations:** There's a loose community exploring how the interactive medium enabled by computers can be used to communicate and think in ways previously impossible. These ideas start, as many ideas in computing do, with work done by Douglas Engelbart [9] and Alan Kay [10].

  More recently, Explorable Explanations have started to reimagine what an essay can be in this new medium. This started with Bret Victor's [11] and has been further developed by amazing examples (eg. [12, 13, 14]).

  There are also explorations of how we can augment our ability to think in this new medium, bringing previously inaccessible ideas within reach. Once again, Bret Victor has wonderful ideas [15, 16], as does Michael Nielsen [17, 18].

- **Research Distribution:** Over last few decades, there's been a big push for research to be freely available online. This includes the formation of arXiv.org and PLOS, and journal editorial boards resigning to start open-access journals.
  Increasingly, the challenge is filtering accessible content. Karpathy's ArXiv Sanity is a lovely tool for this. Crowd curation by online communities also helps a great deal.

- **Open-Notebook Science:** Open notebook science, like Dror Bar-Natan's Academic Pensieve, and massively-collaborative research projects like the Polymath Project, separate the sharing of results from formal publishing.
  This seems really important. Traditionally, if one doesn't turn research into a paper, it's basically as though you didn't do it. This creates a strong incentive for all research to be dressed up as an important paper, increasing noise.

- **Discussion of Debt and Distillation:** A number of mathematicians have discussed what we call research distillation. Some great comments and references are collected in this MathOverflow thread — I'd draw particular attention to Thurston's account of accidentally killing a field by drowning it in research debt in section 6 of [19].
  Other nice thoughts in this space include the idea of an "open exposition problem" [20], Grothendieck's "rising sea" approach to mathematics [21], and recent calls to value conceptual progress in CS more [22].

Author Contributions

This essay was primarily written by Chris Olah and illustrated by Shan Carter.

Footnotes

1. That is, really outstanding tutorials, reviews, textbooks, and so on. [↵]

2. Because most work is explained poorly, it takes a lot of energy to understand each piece of work. For many papers, one wants a simple one sentence explanation of it, but needs to fight with it to get that sentence. Because the simplest way to get the attention of interested parties is to get everyone's attention, we get flooded with work. Because we incentivize people being "prolific," we get flooded with a lot of work... [↵]

3. More formally, if *N* people are trying to understand each other, it takes each one $O(1)$ effort to write an explanation of their ideas but $O(N)$ effort to understand each of the other *N-1* people's ideas. The result is that energy cost looks like $O(a + bN)$ where *a* and *b* are coefficients for the trade off between energy on the explanation side and energy on the understanding side. That is a is the energy spent on explaining and b is the corresponding effort needed to understand. This is similar to ideas in [4] [↵]

4. For example, Christopher's average blog post is read by over 100,000 people; if he can save each reader just one second, he's saved humanity 30 hours. [↵]

5. Often, some individuals have a much more developed version of an idea than is publicly shared. There are a lot of reasons for not sharing it (in particular, they're often not traditionally publishable). [↵]

6. There are a lot of perverse incentives that push against explaining things well, sharing data, and so forth. This is especially true when the work you are doing isn't that interesting or isn't reproducible and you want to obscure that. Or if you have a lot of competitors and don't want them to catch up.
However, our experience is that most good researchers don't seem that motivated by these kind of factors. Instead, the main issue is that it isn't worthwhile for them to divert energy from pursuing results to distill things. Perhaps things are different in other fields, or I'm not cynical enough. [↵]

7. There is a strange kind of informal support for people working on research distillation. Christopher has personally benefitted a great deal from this. But it's unreliable and not widely advertised, which makes it hard to build a career on. [↵]

References

1. **The Tau Manifesto**   [link]
   Hartl, M., 2013.

2. **Interpretive Labor**   [link]
   Van Nostrand, E., 2015.

3. **The utopia of rules: On technology, stupidity, and the secret joys of bureaucracy**
   Graeber, D., 2015. Melville House.

4. **The Mythical Man-Month**
   Brooks, F., 1986. Tutorial, Vol 11, pp. 35--42.

5. **Visual Complex Analysis**
   Needham, T., 1998. Oxford University Press.

6. **Visual Group Theory**
   Carter, N., 2009. MAA.

7. **A visual explanation of Jensen's inequality**

   Needham, T., 1993. The American mathematical monthly, Vol 100(8), pp. 768--771. JSTOR. DOI: 10.2307/2324783

8. **Visual Differential Geometry**

   Needham, T., Work In Progress.

9. **A Research Center for Augmenting Human Intellect**   [link]

   Engelbart, D.C. and English, W.K., 1968. Proceedings of the December 9-11, 1968, Fall Joint Computer Conference, Part I, pp. 395--410. ACM. DOI: 10.1145/1476589.1476645

10. **Personal Dynamic Media**   [link]

    Kay, A. and Goldberg, A., 1977. Computer, Vol 10(3), pp. 31--41. IEEE Computer Society Press. DOI: 10.1109/C-M.1977.217672

11. **Explorable explanations**   [link]

    Victor, B., 2011.

12. **Visualizing Algorithms**   [link]

    Bostock, M., 2014.

13. **How to Fold a Julia Fractal**   [link]

    Wittens, S., 2013.

14. **Back to the Future of Handwriting Recognition**   [link]

    Schaedler, J., 2016.

15. **Media for thinking the unthinkable**   [link]

    Victor, B., 2013.

16. **Learnable programming**   [link]

    Victor, B., 2012.

17. **Thought as a Technology**   [HTML]

    Nielsen, M., 2016.

18. **Toward an Exploratory Medium for Mathematics**   [HTML]

    Nielsen, M., 2016.

19. **On Proof and Progress in Mathematics**   [PDF]

    Thurston, W.P., 1998. New directions in the philosophy of mathematics, pp. 337--55.

20. **A beginner's guide to forcing**   [PDF]

    Chow, T.Y., 2009. Communicating mathematics, Vol 479, pp. 25--40. American Mathematical Soc.

21. **Recoltes et Semailles**

    Grothendieck, A., 1985. Reflexions et temoignage sur un passe de mathematicien, pp. 621.

22. **Statement on conceptual contributions in theory**   [link]

    Aaronson, S., Borodin, A., Chazelle, B., Goldreich, O., Goldwasser, S., Karp, R., Kearns, M., Papadimitriou, C., Sudan, M. and Vadhan, S., 2008. Shtetl-Optimized.

Updates and Corrections

View all changes to this article since it was first published. If you see mistakes or want to suggest changes, please create an issue on GitHub.

Reuse

Diagrams and text are licensed under Creative Commons Attribution CC-BY 4.0 with the source available on GitHub, unless noted otherwise. The figures that have been reused from other sources don't fall under this license and can be recognized by a note in their caption: "Figure from ...".

Citation

For attribution in academic contexts, please cite this work as

```
Olah & Carter, "Research Debt", Distill, 2017.
```

BibTeX citation

```
@article{olah2017research,
  author = {Olah, Chris and Carter, Shan},
  title = {Research Debt},
  journal = {Distill},
  year = {2017},
  note = {https://distill.pub/2017/research-debt},
  doi = {10.23915/distill.00005}
}
```