

## Our New Job Description

Annette Thomas  
*Macmillan*, [annette@macmillan.co.uk](mailto:annette@macmillan.co.uk)

Follow this and additional works at: <https://docs.lib.purdue.edu/charleston>



Part of the [Library and Information Science Commons](#)

An indexed, print copy of the Proceedings is also available for purchase at:

<http://www.thepress.purdue.edu/series/charleston>.

You may also be interested in the new series, Charleston Insights in Library, Archival, and Information Sciences. Find out more at: <http://www.thepress.purdue.edu/series/charleston-insights-library-archival-and-information-sciences>.

---

Annette Thomas, "Our New Job Description" (2012). *Proceedings of the Charleston Library Conference*.  
<http://dx.doi.org/10.5703/1288284315072>

## Our New Job Description

*Annette Thomas, CEO, Macmillan*

*The following is a transcription of a live presentation at the 2012 Charleston Conference on November 8, 2012. Video of the session is available on the Charleston Conference website at [http://katina.info/conference/video\\_2012\\_thomas.php](http://katina.info/conference/video_2012_thomas.php).*

I'd like to start by telling you a little bit about Macmillan because, like all companies, it's completely obvious on the inside and a little less obvious on the outside. Macmillan is an international publishing group. It was founded in 1843 by Daniel and Alexander Macmillan in Scotland. In 1869, they founded Nature, and the scientific journal Nature, actually came out of the tobacco salons of the 1850s and '60s in London where intellectuals would gather to discuss the latest scientific and literary topics of the day. So, it was the Macmillan family that actually founded Nature in 1869. Since then, obviously, there's been a huge amount of development for Macmillan, and we are a company that spans, we say, two different types of publishing. One you see represented on the slide; that's our science and our education publishing. We also have a consumer book publishing division, which I won't be speaking about today. So, we are science, scholarly, and educational publishers. On the education side, we tend to focus mainly on English language teaching all around the world, particularly in Latin America where we are very, very strong, and in India and China. We also have higher education publishing as well here in the US and around the world. Today, my talk is going to focus quite a bit on our scientific publishing out of Nature Publishing Group and Digital Science. We also have a very vibrant humanities and social science publishing unit, Palgrave Macmillan, and I will refer to that just briefly.

Before I go into the main part of my talk today, I've been asked to tell you a little bit about myself, which is not something I usually do, so please do bear with me. I joined Macmillan 19 years ago, now almost 20 years ago, after finishing my PhD. I did an undergraduate at Harvard in biochem and biophys, and I did my graduate work at Yale studying the cell biology of a neuron. Most of you

know that graduate programs in the states tend to be quite lengthy affairs, and mine was no exception. My PhD took me six years to finish at Yale, and at about year number two, I decided this was great fun, and we were actually having great fun, and we were publishing well, including publishing in Nature, but I decided that I would try and find something that I could do with my research background that wouldn't necessarily be research, and what I have realized over the years is that I think I made the right decision, but being quite a young person then, I probably made the right decision for all of the wrong reasons, as often happens. But I finished my PhD and at the end of that, it was 1993, I had the opportunity to join Nature as an editor. So, just to give you a bit of context, 1993 was just the beginning of digital, insofar as e-mails and servers and things like that, well it was for most companies. It wasn't for Nature Publishing Group. It was also, I think, when scientific publishing was really starting to enter into a new phase; it was becoming more competitive, even for publications like Nature. So, I joined them in London as an assistant editor, so my job was to select the manuscripts and see them through peer review in the areas of cell biology and neuroscience, and I did that for a couple of years, and it was a fantastic job, particularly for someone just finishing up their graduate work. We were exposed to a lot of different science, a lot of different scientists— incredibly exciting. And in the years that followed, I had a number of different roles at Nature. I launched Nature Cell Biology. I was the editor of that title when we launched that in 1999. I held some commercial positions in between. I was the publisher of the Review journals which we launched in 2000, Nature Reviews. We now have several titles under that banner. So, we launched the Review journals in 2000, and just after that I was appointed the managing director of Nature Publishing Group, and that's really when our phase of digital development began, setting up our Nature.com platform, site license program, the many different business models we tried. I'm sure there are people in the room that have as

fond memories of that time as I do. And over that period, then we started to expand more rapidly the Nature brand into new research journals, and our academic publishing became a bigger part of what we were doing. And we had several different innovations that we were creating through our web team which has actually evolved into Digital Science. We will talk a bit about that, but in 2007, I was then appointed Chief Executive of Macmillan, which is the parent company of Nature Publishing Group, and I have held that role for five years. So, I stand here as the Chief Exec. of a medium-sized publishing company. We are still family-owned. In '95 the Macmillan family decided to exit publishing and they found an excellent home for Macmillan. They were looking for a home that would keep Macmillan "Macmillan," and not separate it up into its component parts, and they found that home with the Holtzbrinck Media Group, which is based out of Germany, and we've been part of Holtzbrinck since 1995, and that has indeed been a very good home for Macmillan. We are still family-owned, and I say that because it gives us a certain perspective on what we do, and I think you'll see some of that come out in the rest of my talk.

We're very, very focused on science education; it's the main area of focus for us. We take a very long-term view of what we're doing and trying to evolve our business, because we feel that we have an obligation to all of our customers: our researchers, our scholars, our professors, our teachers, our students, to evolve over the long term and not look for short-term solutions to long-term issues. And that's a little bit about our company; it's a little bit about me. Scientific background, scientist at heart, I still have not managed to find the time to go off and do an MBA or a mini-MBA, so I have a lot of people who support me in a lot of the different things that I do, and it's my pleasure to now move into the main part of my talk where I will be able to share with you our perspective on how publishing's evolving and a little bit of the hard work that they've been doing over the last couple of years.

So, I'm going to talk quite a bit about science today, and when we think about how publishing is evolving, we think first and foremost about the

scientist, and I know that sounds really obvious, but many publishers operate in what we would call a "B2B" kind of market, and that means they are very focused on, not necessarily the end customer, but the products and how those products are ultimately going to go into the market. We try and take a really customer-oriented approach and always remember that we are here to try and make scientists and the scientific research process more effective and more efficient, and when you put on that hat, when you take that perspective, what we can do as publishers actually expands far beyond what we would call traditionally publishing discoveries. Of course, that is an important part of the scientific process. It is very important to scientists to be able to communicate the findings, record their findings, and get credit for their findings; but particularly now with the evolution of digital technologies, there are many, many more ways that we can think about how we can help make that scientific research process more effective, more efficient, more successful.

Learning about discoveries is the process the scientists increasingly have to grapple with and spend a lot of their time on. What's important to me as a scientist in my particular field? Who is important to me? Why is it important to me? How do I keep up to date with the increasing deluge of information that's coming at me from all different angles? How do I plan my experiments? I may know what I want to do, but how do I know the best way to do it? What's going to give me the best chance of success?

Actually doing the experiments is also increasingly, particularly in labs that are now much, much more interdisciplinary than they used to be, quite a challenge for scientists. Evaluating the results, sharing the results; sharing the results is not just about publishing your results. I was thinking about who in my community and beyond will be interested in these results and how do I best get them that information. At the very top of the slide you can see there is a slightly lighter circle which is really more about the scientists themselves. So, after I publish, how do I gain and grow and enhance my reputation? How do I obtain funding, obtain collaborators, so that I can

grow my lab, so I can do more experiments, and continue the cycle of success? When we think about how we want to evolve what we do, we very much have this picture in mind and we call them “pain points.” What are the pain points the scientists have that are keeping them from being as successful as they would like to be, and what can we do, not necessarily on our own, often in collaboration with others, to address those pain points?

I'm going to share with you a few of the projects that we're working on, and there's other groups out there, some attending the conference over the next couple of days, that also working on some of these problems. So by no means do I think we have all the answers, but we're trying some approaches that I hope you'll find of interest.

So, publishing discoveries, you might say “there's nothing new here.” We've been publishing discoveries since science began, certainly even before Nature was launched. But publishing discoveries has changed. We all know that business models, new business models, have been introduced, open access, author processing charges, Green OA, Gold OA, there's a lot of new types of business models, and Nature Publishing Group is also embracing that. Nature Communications was a journal that we launched some 2+ years ago. It has a mixed business model with subscriptions but also with open access. Scientific Reports was launched more recently. That's a purely open access journal with author and funder charges.

But the business models are not the only thing that's changing. It's also how scientific research is actually communicated and published that's changing. In 2001, Nature published the Human Genome. That was a seminal, seminal research paper. Three billion nucleotides sequenced. I didn't bring it with me because I had only carry-on luggage knowing the transportation challenges of getting from London to Charleston with the storm and everything else, but had I brought it with me, I would be showing you now “2001: The Human Genome,” a telephone book-like directory which recorded the three billion nucleotides that were sequenced by the Human Genome Project. Last

month, Nature and two other publishers published “The End Code Project.” Now, this is really interesting because of those three billion nucleotides, only 1.2% encoded genes, so what are the other nucleotides doing? In 2003, the End Code Project was kicked off to answer that question. What are these gene deserts for? Do they do anything? Some of you might be familiar with the term junk DNA. Well, is it really junk DNA?

What we found out last month in the publication of the End Code Project is that it's not junk DNA. Eighty percent of that DNA is actually encoding regulatory elements, promoter elements, and it's there that our uniqueness comes. We have more diversity in that so-called “junk DNA” that's regulating our genes than we do in our genes ourselves. This is a really important discovery. Now that we know that 70% of the nucleotides have these important functions, we just have to figure out what they actually regulate.

But it's not just science that has moved on. Also, how it's published has moved on. “The End Code Project” is not a telephone book directory of nucleotides. “The End Code Project” was published last month completely open, in collaboration between three publishers, and in this unique format that you see on the slide at the bottom, the bottom left in the black, it sort of comes out where there's these threads where we've connected the main concepts the scientist will want to know about this project; he can click on each bubble and the threads come up leading the scientists to the relevant papers. Once he goes into those research papers, much of the data can be actively mined and interrogated on the spot by the readers. This is a very different approach than in 2001, so, it's not just business models that are changing; it's how publishing is actually changing. And whilst today's talk is mainly about science and scientific research, it's also changing in the humanities and social sciences.

In Palgrave Macmillan we announced earlier this year Palgrave Pivot, and on the surface this seems like a very straightforward concept, but within HSS it's been warmly embraced. What is Palgrave Pivot? Humanities and social science scholars have two types of formats, and you all will know

this. They have the journal format, short format, and they have the monograph too, long format. But that's a very sort of physical world-type constraint. Palgrave Pivot brakes open those boundaries, allowing the scholars to publish at whatever length suits their research, but it also introduces new and flexible business models into the HSS space which have yet to really take hold. So, there's innovation and publishing around business models, around how publishing is happening, around mining data. This isn't yesterday's type of publishing. It's very much today and into the future.

Learning about discoveries: again, the information overload problem. Traditionally, scientists would have turned to content such as Nature Reviews and Scientific American, also in our stable, and they still do. These types of seminal pieces of work provide really useful overviews for readers. But there is a new type of way to keep up-to-date, and before I go into this just to refer to some experiences I had recently.

Just last week, I spent the week in the Bay Area in San Francisco, and we visited UC Davis, UC Berkeley, UCSF, and Stanford over three or four days. We met with graduate students. We met with post docs. We met with PI's. We met with librarians. We met with professors, teachers, and students, talking about all manners of issues around scientific research, around peer review, around publication, business models, Prop 30, which did pass in the election, which is great news; a lot of different issues in a very challenging market.

For me it was one of the most interesting visits that I've had in some time, particularly around the graduate students, and listening to them, how they keep up-to-date with information. It's certainly not how I kept up-to-date with information when I was in grad school. They are mining the social web. They are mining Facebook, Twitter, and social media. They are interacting with their colleagues. They are using, yes, traditional table of content alerts as well, but it's a much richer way of mining information and keeping up-to-date. And my belief is that, going forward, well, there's not going to be one way, there's no magic bullet that will keep you up-to-

date as a scientific researcher, but it will be a combination of tools that will be used and that will involve technology and algorithms and search. It will involve user-generated content as well, and ultimately, I think, in some cases it will also involve that editorial overlay that we are so familiar with, that quality control. We're developing several tools in this area, and I just have a couple to tell you about today.

So, altmetrics: there is an altmetric movement afoot, and by altmetric I mean alternative ways to measure impact, to measure relevance. There is a disbursed conversation going on online as we speak about scientific research. In the social web, right now articles are being discussed and being debated all over the place. But how do you track that? If it's your research, you're interested. If it's research that you should know about, you're also interested, and you'll want to join the conversation. Altmetric is a tool that allows you to do that. It tracks article research, articles being mentioned in Facebook, in Twitter, in blogs, but also in the mainstream media: BBC, New York Times, CNN, Scientific American; and it allows scientists to track the research they're most interested in, and obviously join in on that conversation should they wish to do so. Increasingly this is an important measure of impact. Impact isn't something necessarily that's just cited and measured two years later, the kind of impact we're all familiar with. It's the impact we're having right now. Any of us in this room could be having that type of impact, and Altmetric is a tool that allows scientists to track that and to participate in that. It tracks over two and a half thousand sources. It's three million papers and growing and has already over five million social media mentions.

So, ReadCube is another tool that we've developed, and ReadCube does a couple of things. It helps scientists organize their PDF collection. I mean, did you ever stop to think that scientists have better tools to organize their music and their photos than they do their PDFs? It's quite something. But not only does it allow a scientist to organize their PDFs, it also allows them to annotate and find it again with a powerful search tool. It also, really importantly, allows them to

discover new research, because ReadCube is a tool that the more you use it, the smarter it gets, and it will send you personalized alerts on the content, not that you already know about, but the content that should be of interest to you based on what you're already storing and what you're reading. I have to just a short video here too (Video available online at <http://www.readcube.com/#features>, accessed 1/14/13).

So, those of you with particularly good hearing, perfect pitch, will realize that that was my voice doing the voice over, so, you know, we all muck in together at Macmillan. That's a little bit about ReadCube, but it goes a bit beyond that. In the last month we started a pilot with Rick Anderson and his colleagues at University of Utah to use ReadCube to allow their patrons to access content that they don't already have site license subscriptions to, and that means that the researchers at University of Utah can go in through ReadCube and purchase at very, very affordable prices, not just access to the content, but access to the rich PDF format of ReadCube, which you just saw, and they can rent that content by the day or longer. They can read it online; the prices are very, very affordable, lower than the usual one-time access fees. The only thing they can't do with it at the moment is to print it or to share it, so it's a different type of access business model which makes it more affordable and richer. Off the back of this pilot today, actually, my colleagues here in Charleston are announcing that all of the Nature content now will be available through ReadCube for purchase and access in this multiple of ways and using the rich ReadCube format. I think they're hosting the party tonight at six, so if any of you are interested in learning more about ReadCube, do take a look and take it up with them directly, because they are the experts.

So, planning experiments. This is what we might generously call an information rich environment. Before I was a scientist, when I was still a young, young, young person and reading Scientific American, I had this image of what a lab would look like, and this wasn't it, but this was very much what the lab that I worked in looked like, and I can tell you, I was visiting labs last week, and

things haven't really changed. Scientists need help in order to organize their laboratory environment. We're talking about consumable samples, data sets, shared equipment, and protocols. How would you go about organizing your lab? How do you organize your stuff? And by the way, when that postdoc leaves, and six months later he's gone, you need to find out where his protocols, where his samples, where his consumables were because you're going to need to be able to build on and reproduce those experiments. How do you do that in a lab? Labguru is one tool that we've developed that helps scientists organize their stuff, and if they can organize their stuff better, they can spend less time on that and more time making the big discoveries.

1DegreeBio takes a different approach. It's essentially a marketplace for consumables in science where scientists can rate and rank and share their experiences. In the first instance it started around antibodies, because antibodies are a huge part of much of the molecular biology research that is done, but it is extending now beyond that to include all aspects of consumables in labs. These are two different ways of trying to really help scientists record what they're doing and organize themselves, but also be more effective at planning their experiments in the first place.

What about doing experiments? Okay, so compliance is not the sexiest topic in the world, but increasingly when you're on campus, campuses have to be run in a professional way. Compliance is a part of all of our lives. But compliance in a lab, if you think compliance in a lab 20 years ago versus compliance now, it's completely different. BioRAFT is a tool that helps scientists, and particularly helps departments, organize their compliance programs: their training, their paperwork, that they have a record that they are actually setting themselves up to do the experiments in a proper way.

Okay, sharing results. So sharing results is not, these days, just about publishing a paper. Scientists do a lot of research, and a lot of it never, ever gets published; and even research that does get published, if you publish genes, or if you publish proteins, then it is very obvious where

you should store your data. But most of the research that is published, and most of the research that's unpublished, there's nowhere to store that data. Why would you want to store the data? Well, you would want to store the data because you want to retrieve it at some point, but actually you want to store that data and make it available because you want to share it, you want to get credit for the work that you've done, the work that you've done that doesn't have a natural home in one of the big existing databases, or the work that you've done that's perhaps had a negative result, and most science is negative results.

Figshare is a tool that's been developed that allows scientists to do just that. They can store their data, they can make it available to others, it's citable, and it's trackable, so that they can get credit. Some of the statistics on Figshare—at the moment, it has thousands and thousands of different types of data sets that have been downloaded, and the usage of Figshare now extends to countries all over the world and some of the top labs all the way to some of the smaller labs. We're trying to solve the problem of what you do with data that normally would sit in your cupboard or sit on your computer, making that discoverable, making that trackable, making it citable.

So, that's a little bit about some of the things that we are working on at Macmillan, at Nature Publishing Group, and Digital Science, in particular. I'm often asked, in fact, we have a lot of debate in our company, "What makes a service or a solution indispensable? What makes it really indispensable? When you look at that map what makes something really a must-have?" I have four characteristics that I look for, and not every tool or service or solution that we think of ticks all of these boxes, but these are things that I think are important.

First of all, it has to demonstrably and measurably improve an outcome. Not just because I say it does or the marketing material says that it does, but we have to be able to measure that improved outcome. Second, it should be global and set a global standard. If we can introduce more global standards, for example, around the way that we

store data, it makes it much easier for others to do the same, but also to build applications off the back of it. Third, I'm doing this now from memory, it should contribute to and benefit from network effects, so the more that that tool is used, the smarter it becomes; the more valuable it is for me as the user, the more valuable it becomes for all the users. And I'm thinking of number four, which will come to me.

When we look at this map, I'm often asked "What is the business model?" Now this was the topic of conversation last week. "What is the business model?" Well, the fact is that the business model is different for different parts of this particular landscape, and what I look for, what we look for as a group, is we look for the entire ecosystem to be commercially sustainable in the long term. Some of the things that we focus on and we produce, we do that because it generates a network effect. It draws users in; it's valuable to them. Other things that we produce actually retain our users within the network and other things that we do, in and of themselves, will generate revenue. But it's that rich ecosystem that we are looking to establish that needs to be commercially sustainable for the long term. Why does it need to be commercially sustainable for the long term? Well most importantly, for me, it needs to be commercially sustainable because we want to be able to continue to invest in the future, and also as a medium-sized privately held company, we want to be certain that we always have that bandwidth to keep investing for the future. And so for us it's the mix. It's the mix of everything that we do together that has to be commercially sustainable. That gives us a huge amount of flexibility, I think, and freedom to experiment, to think creatively and innovatively about how we're going to pull these various services together. So there's not one business model; it's not one-size-fits-all. It really depends on how the various services that we produce interact together.

So, going back to my four indispensables, I said that it should be global, and set global standards, that it should have demonstrably measurable outcomes, that it should create and benefit network effects and now I'm just trying to think of

the last one, and I'm jetlagged. Apparently I'm going to be interviewed later this afternoon, so for any of you who are really, truly interested. Now I'll get to measure your interest in my talk, if you really want to know what number four is, because that's the kicker I'm sure, that's the magic dust, then by all means do follow it up in the interview.

It has been an enormous pleasure for me to share with you a little bit with you of what we're doing at Macmillan. We do not have all the answers, and we don't operate on our own within our company, but we have many collaborators in the library community, in the scientific community, with other companies. I'm not presenting this as the answer. It's going to be a very rich answer with contributions coming from a lot of different places, I'm sure. But this is a little bit about what we're doing and a little bit about our approach.

Last week, when I was at UCSF, the Provost said to me "Ah, publishers. I feel sorry for you guys," and we then went on to have a conversation about some of the issues he was interested in, some of the challenges he was interested in, and at the end of it he realized that he's interested in exactly what we're interested in.

I'm completely optimistic for the future of publishing, and I'm optimistic because there's so

much more to do now than there was 20 years ago. Twenty years ago we spent most of our time on the publishing discoveries, and whilst that will always remain important, there is now a very rich landscape that digital technology allows us to really, truly think innovatively and creatively about. So I am very optimistic. And in a funny way, I can see that the librarian community and what you all do is very much aligned with this as well. We both have the same goal, which is to make scientists and the scientific research process more effective, more efficient, more successful. But we're both at a much, much richer landscape, and that provides a lot of challenges but also a lot of opportunities for how we might go about doing that, and I very much hope that we can work on it together in the months and years to come because this is an exciting place to be.

There is a huge amount of opportunity here. It's not easy, but it's definitely there. So Accentuate the Positive, I believe, is the theme of the conference. I believe there is a huge amount of positive, absolutely, and anyone who is coming into publishing today, or anyone that is coming into information sciences today, really has the world as their oyster. Thank you very much for inviting me.