Metadata Is Not A Thing

Before I delve into issues around metadata and workflow, I want to back up somewhat from our offices, our conferences, our publishing houses.

Away from Midtown Manhattan. Away from New York City entirely.
Away from North America.

In fact, let's have a look at the world as a whole. Because that's why we're here, isn't it?

How many of us in this room got into publishing so we could send emails back and forth to editors and production people and agents and finance and contracts people? When thinking about what we wanted to do with our lives, did any of us say, "I want to spend my days sending urgent emails to Amazon to fix the typo in the description on that book?"
I think we're made of better stuff than that. And I certainly think we're capable of seeing the meaning in what we do.

We're in publishing. As Margaret Atwood reminded us last year, we are doing something more than just ordering paper from mills and binding pages together.

And there are lots of different ways to publish.

Basically, as publishers, we GET A MESSAGE OUT.

That message could be a lot of things - and if the stars align, we manage to make a living actually believing in the message we're trying to get out - but what's important to this particular conversation is the spreading of that message.

That's what we do. We spread the word. That's the business we're in. We are professional word-spreaders.
We want to preach the word – be it how to cook a chicken, a story that’s really gripping, a method of meditation, a new interpretation of Hamlet, a biotech discovery, an economic perspective – whatever the word is, it’s our job as publishers to spread it.

So who are we spreading the word to?

We really don't want to be preaching to the choir. The Twitter echo chamber is a very finite world. They’ve heard it all before. We want to be spreading the word
to people

who are hungry
to hear it.

We're talking to everyone. People all over the globe, in all professions, of all ages. Our mission, our purpose, is to reach everyone.
This is hard. But it’s what we’ve been tasked with, just by being in this industry.

And I’m going to tell you how to do it.
If you’re an acquisitions editor, you know the author, the title, the formats the book will be published in, what the book’s about. And it’s not just WHAT you know, but WHEN you know it - that dictates how the message to the audience is going to be refined and changed. The title could change. You could add an author or have a noted expert write an introduction. The focus of the book could shift. This is all critical information for readers to know.

If you’re a managing editor, you know the ISBNs of the formats, the dates the book will be published, the prices of the editions, how many will be in the first print run, when the Kindle conversion is coming back in-house. This information is highly volatile - it changes a LOT. And it is all critical for readers and the supply chain to know. You may know that there will be 4 formats when you create the first schedule - then you may later know that there will actually be 5 or 6 formats because you are selling to some new vendors who need their own proprietary files.

If you’re a marketer, you know how the book is supposed to be sold – what categories and keywords are appropriate, what message needs to be framed in the descriptions and flap copy, which excerpt is most gripping. And this, too, could change - depending on editorial influence, or production needs.

You know where I’m going with this.
Increasingly, as we leave the physical realm during this period of digital upheaval, metadata for books is more critical than it ever was.

As Brian O’Leary so deftly pointed out at last year’s TOC, when publishers liberate their content from its containers, metadata is pretty much the only indication that the content exists, and definitely the only indication as to where it can be found.

And literally everybody in this room has a hand in that.
In each step of the publishing process, information about the book - metadata - gets created. It doesn't always get used – sometimes it gets forgotten. Sometimes it gets stashed away on an intern's hard drive...
...and after you’ve fired him, it’s never seen again. Because he password protected his hard drive with that one Skrillex lyric you’ve never been able to figure out because that would require you to have to listen to Skrillex. Which is impossible, really, when you think about it.

Your communication with the people who are hungry for the word you’re spreading...should not be trusted to this guy. That information is too valuable, and it’s changing too fast.
It helps to keep that metadata in a single system that everyone has access to. (With varying permissions, of course.) This way, when you fire your drunk intern, you can still get to the metadata you’ve asked him to enter (and correct it or enhance it or change it). You know EXACTLY which titles are in your spring list – because there are not six spreadsheets floating around with different ISBNs on them. My boss, Fran
Toolan, calls this system a company’s “single source of truth”. My colleague Rob Stevens calls it “one database to rule them all”.

Let’s look at this functional map again:

![Metadata as a Functional Map](image)

Notice how each division of the publishing house has something to say about a book. Everyone participates. Everyone is accountable to everyone else.

If Production screws up a piece of information, Editorial’s going to know, and so is Acquisition, and so is Marketing – because those folks are going into the same database to put in their information, and to use what’s already been laid down.

By centralizing the product database into a single hub, each function across a publishing house can do their work – and in the process, vet the metadata because they have to use it to do that work. The metadata becomes clean as the book moves through its development cycle... because it is exposed to so many people whose jobs are so very different - all of whom have to use the metadata in different ways. But all of those people are working towards the same goal – getting the word out about that book – in some cases, just getting the word out to each other, but ultimately the greater purpose of getting the word out to the reader gets accomplished as a byproduct of all this communication to a central hub.
It’s the hive mind at work.

Or, if you have a more sinister point of view....

So when I say metadata is not a thing, what am I talking about?
As metadata has taken off as a subject in itself, I’ve started thinking about metadata and communication in terms of audiences. Just as I am gearing this talk to a certain audience – and how I’m communicating (and even what I’m communicating) is different for this audience than it was the time I gave a very different version of this at NISO’s ebook conference, or different from the phone conversations I have every day with my clients – communication among anyone has to take audience into account.

For the metadata you’re entering or sending out, who’s the audience? Who are you sharing this information with? What works for you and your colleagues and your trading partners?
Editors are trying to communicate one set of information; production has another whole set of priorities to worry about. There’s metadata that, while very valuable internally, publishers do not want going out the door – such as budgets, sales figures – and metadata that publishers very much want to go out the door, such as good reviews.

As you’re communicating – as you’re part of the hive mind – you have to think about who’s receiving that communication. Do they really need it? Are they missing
something they should probably have, something that would help spread the word about the book in some way? (Carton quantity and case pack indicators are a huge help.)

It helps to put yourself in the shoes of the people you want to communicate with. You have to speak their language.

For example: Amazon's not putting inventory information up on its website for customers. They put up a notice of availability.
but not actual stock numbers - even though they receive that information from publishers and distributors. That’s information strictly for order management. (And I need to thank @mikecane for the above image - he was very helpful.)

Nor do they put up the price at which they acquired a book – again, that’s not the reader’s business and they don’t want to communicate that to the wider world.

But it’s the sort of information that goes back and forth between Amazon and publishers and distributors all the time.

So readers

require different information than booksellers
which require different information from distributors
which require different information from libraries.

So metadata’s not a crystallized thing, an object like a piece of glass. It’s much more fluid than that - like particles of sand themselves that blow around - it’s a somewhat unruly form of communication.

Having a “metadata manager” or a “metadata czar” at your publishing house to manage the object that is metadata... is great in theory – but given all the different points of communication along the publishing chain, it’s tough to put one person in charge of what, to the untrained, can seem like a logistical nightmare.
I grew up near the beaches in Southern Delaware and in the summer, every Sunday after church without fail, my parents would get the 3 of us kids into the car and haul us over to Bethany Beach for the day. We would swim and dig and hunt sand crabs, and we would have dinner on the sand, and eventually get back to the house all sunburnt and...covered in sand. Technically, it being the 1970s, it was my mom's job to ensure household cleanliness. She was the Sand Manager. But of course, the sand was everywhere...

and it wasn't until we were older and could each manage our own sand that the house managed to stay reasonably clean on Sunday nights.

Equipped with this bit of wisdom, I would argue that instead of having a Sand Manager, each function
should be in charge of what it's communicating, to whom it's communicating it, and how it get communicated. I'm arguing that instead of a single metadata guru, that responsibility needs to be distributed across the organization - just as publishers are looking at agile software development for cues to new ways of publishing, we need to be thinking about agile metadata development. That, like any good hive mind, each function inside a publishing house bears equal responsibility and equal expertise for a very large, very complex undertaking.

In turning metadata into a thing - a project, a function in itself, something isolated - publishers risk marginalizing it.
And not paying attention to it - it’s the “metadata manager's job”, after all.

So whatever our individual job actually is – our ultimate goal, the reason why we got into this line of work in the first place, is to get as many books into the hands of as many readers as we possibly can.

THAT is why we are all here – that is what each of us do, in our own way, every day.
Metadata, communication about our books, gives us the ability to get as close to that goal as we have ever been in the history of publishing.

Filling out those fields in a database seems like an extra step, yet more work being piled upon us, and God knows it’s about as thrilling as watching paint dry and not immediately creative or interesting - but doing that work is what will ensure that we can do what we came into this business to do.
Which is to get a book into the hands of everybody who wants one.