Project Process

T. Christian Miller
ProPublica
t.christian.miller@propublica.org
@txtianmiller
http://tinyurl.com/IRE16-Write

Step 1: Digitize

I can't emphasize this enough. These days, I try to digitize every single piece of a story—notes, photos, audio, video, documents. How? A couple of useful tools:

DocumentCloud – Allows you to upload PDFs and performs Optical Character Recognition, aka OCR. OCR takes a document and scans it to make it searchable by word. Not perfect by a long shot. But a start. DocumentCloud also has amazing embed features allowing you to highlight and annotate passages. And, finally, it has a janky but usable timeline creator and a very good entity recognition engine. Available to IRE members.

Google Docs – Does much of the same thing as DocumentCloud and often integrates better into newsrooms that are using Google suite. Also automatically OCRs documents.

Transcription – Overseas transcription services have brought the cost of transcription down to a buck a minute, and do a decent job. So if you have an important interview that you want to post online, this is the way to go. Turnaround can be as soon as one day, but will cost significantly more, like $3 a minute. Services like rev.com; Transcription Associates; Transcribe and TranscribeMe all produce decent transcriptions. A warning: if it is a key quote, REVIEW THE TAPE. Just like the NFL.

OCR Scanning – If you have a huge amount of paper documents that you want to turn into searchable PDFs, often the cheapest way is to find a legal services firm in town. They charge around 15 to 25 cents per page, so even if you have hundreds of pages to scan in, the price isn't too dear. Legal services firm are also fast and they have very high quality OCR engines to recognize documents and turn them into searchable text.

Excel or Google Spreadsheets -- I use Excel or Google spreadsheets for almost everything. You don't have to have a computer database project to use Excel. It comes in handy for creating, sorting and organizing even tiny lists of information. For instance, for a story that I did about civilian contractors injured in Iraq, there were just too many cases to keep in my head all at once. So I built a little spreadsheet of the 30 or so cases that I was focused on, and added in bits of data.

Idea Organizers -- If it's a truly long project, you might consider getting special organizing software like Evernote or Microsoft's OneNote, which integrates with Office. Such programs are designed specifically to let you to paste in web pages, keep track of source information and organize your data. I have not used these very much in my own work, but some people love them for their ability to keep everything in one software package.

Step 2: Datafy

Almost every story can benefit from data. Data helps put a story in context. It helps set your story apart from the competition. And it's becoming increasingly easy to do.

A data analysis does not have to be complicated. It can be as simple as writing a murder story and noting the total number of murders that have happened this year as opposed to last. And it can be as complicated as a multiple regression analysis on backdating of option payment at publicly traded companies.

But the point it this: there is almost always data. Don't run from it. Work with it.

Try this exercise. Pull up a random page from a random newspaper. Look at the first couple of stories. Ask yourself: what additional context could be in this story?

How do you find data? Here's a couple of websites at the federal level to get you started. You will find that state and local governments often respond more quickly than the feds.

data.gov – the main repository of data from the federal government. Divided into topics and agencies, it allows you to search data sets that the feds have made publicly available.

fbo.gov – a list of all federal government contracts out to bid. Has useful descriptions of projects and contact names.

USAspending.gov – lists all government contracts and subcontracts that have been awarded. So kind of the follow up to fbo.gov. Keyword searchable, so you can seek contracts in your state or town.

Enigma.io – a fantastic amalgamation of data sets.

Govzilla – this site is designed for corporate intelligence but contains astonishingly useful FOIA information. Essentially, the site continually FOIAs inspection reports for a number of agencies – the FDA, the IRS, the NIH – and makes them available. They have a high cost. But if you need data on deadline, search here.

Step 3: Chronify

No matter what shape your investigative story will take, long or short, narrative or thematic, character driven or topical, there is always going to be a chronology.

The first thing I do when I sit down to report is create a timeline. Here is one particularly crazed example: http://tinyurl.com/ny428c which I used in writing a story on the bombing of a village in northern Colombia called Santo Domingo. It was 11 pages long. But it really helped me see how the operation unfolded. You'll find the story here: http://tinyurl.com/mqf9oz

Three benefits to a timeline:
·   It helps you see relations you might otherwise miss
·   It helps you quickly be able to refer to events
·   You can include the source in your timeline so that you remember where a particular piece of information came from.

I tend to use a spreadsheet for creating timelines. But, cool hint, you can also create timelines in Word, so long as you use a date format like YYYY-MM-DD to begin the paragraph. Word can sort paragraphs by date if they begin with that format. So you can enter information at the bottom of your Word doc, then simply sort to make sure the timeline is in chronological order.

In my humble opinion, timeline tools are still wanting for the reporting side. They are more focused on production than data collection. There are web-based tools like Tiki Toki and Dipity. And there are software versions like timelinejs, from UNC's Knight Lab or TimelineSetter by my own shop, ProPublica. But in one way

or another, I haven't been satisfied by any of them. The spreadsheet or word document works fine.

Step 4: Personify

Now we're getting down to business. You've got to bring the story alive. That means having good characters, who say interesting things.

When taking notes, or talking to someone, I always try to put a couple of asterisks near the quotes that sound good. Then, when I review my notes, I search for the asterisks to create a file just of quotes. I then go through that file and figure out what are the best 10, 15, 20 quotes that I have. Again, two reasons:

- It helps you organize your story. You can begin to imagine transition paragraphs, kicker quotes, opening quotes that will help shape the story.
- You make sure you're getting the best bang for your buck. You're really looking for information said in a pithy, punchy way, sifting through everything that was said to get down to the very best.

The other big thing to look for is characters. This is not, of course, always possible. If you have a story that spans a lot of time with a lot of characters, you may be better off simply telling through chronology and let time be the character. Or there may not be someone who fits the story well – one of the worst things to do is try to 'fit' a character into an anecdote that really does violence to the person.

If on the other hand, you have a combination of a person who speaks in fountains of quotes, whose life story is powerful and moving, and who illustrates your story well, then you have a bit of magic on your hands. Use that person to the fullest extent.

Step 5: Narrafy

This is the really hard part. You've got to figure out how to tell the story.

The good news, however, is that in doing steps one through three, you should have an inkling of what to do.

My favorite story organization structure is the chronology. If you can unspool a story in more or less chronological order, it helps the reader understand what's happening, it makes clear how one event is linked to another, and it makes for easier reading. In fact, I will invite damnation by saying that chronology is almost the ONLY way to tell a story of any length.

Generally, I'll try to write a top to summarize the story and the main points. I usually try to find a quick, compelling scene to which I will return later, or do a simple hard news lead. Then I'll write the nut graphs, some key findings, and a quick set of responses to the findings.

That all should take up about 10 to 15 paragraphs. By then, the reader will have figured out whether it's worth spending his time to dive in. After the top, I hit the brakes, and the story unwinds more or less along the timeline.

My second choice is thematic. In other words, I'll break the story into chunks that explain the issue. I call this the Mixed Bag approach. But even here, I generally try to use anecdotes within the topic areas to develop the story along a timeline.

So that's my story process. Although I have written it in steps, my process proceeds parallel most of the time. Almost from the beginning of reporting, I'm thinking about how to personify and narrafy the story. I'm doing digitizing and datafying. It doesn't happen in a rigid order, but in an iterative fashion, constantly going back and forth when I discover more data, or more characters, or when digitizing data reveals new trends.

It can be a long, tough process. But in the end, I think that readers are seeking powerful, well-told, well sourced stories. We can't always deliver that given the vagaries of deadlines, shrinking resources and disappearing staff. But we can try. And when we get it right, it's awesome.