

THE STATISTICIAN AS PROJECT MANAGER: **STRATEGIES FOR SUCCESS & SERENITY**

Jennifer Thompson, MPH

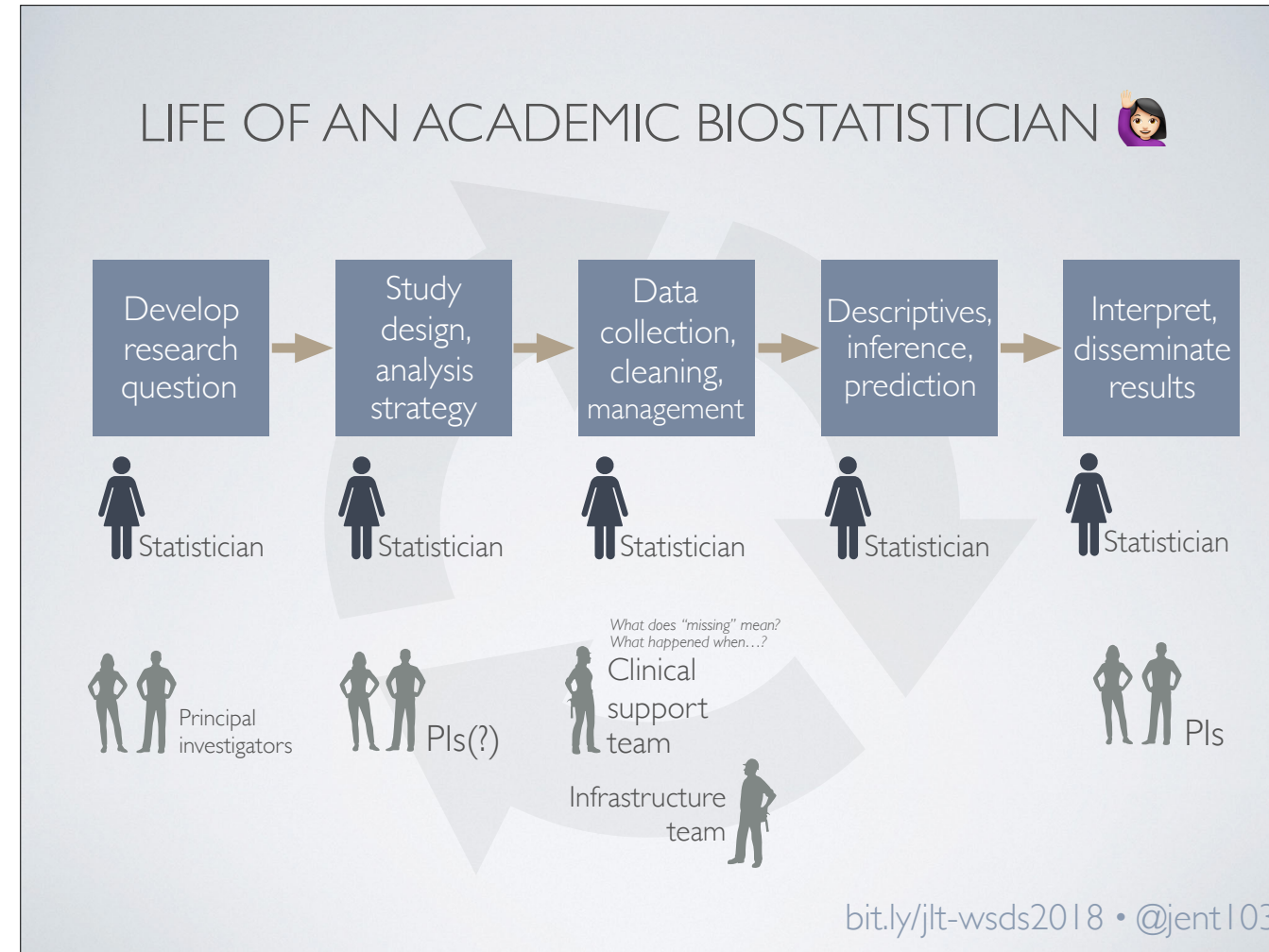
Vanderbilt University Medical Center
Department of Biostatistics
Critical Illness, Brain Dysfunction & Survivorship Center

bit.ly/jlt-wsds2018 • @jent103



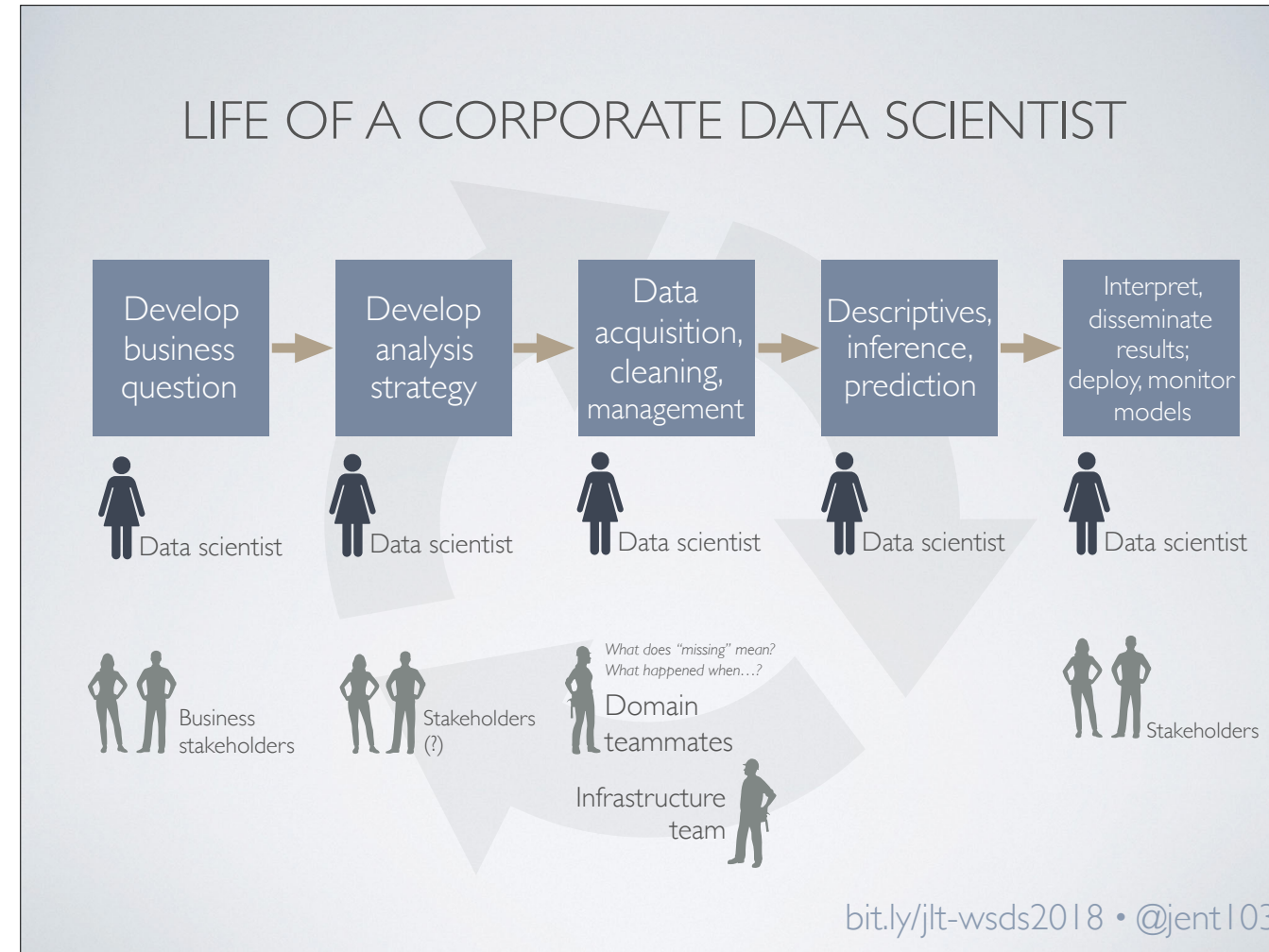
Thank you all for coming! Today I'm going to be talking about how we as "data people" often end up in the role of de facto project manager, and how we can keep our projects on track while remaining calm and moving forward in our technical careers. This photo is of me enjoying wine and cured meats on a hill in Tuscany, and it represents how I want all of us to feel about these tasks at the end of this talk.

For context, I'm currently an academic biostatistician at Vanderbilt University Medical Center, working with a group of over ten principal clinical investigators. That means a lot of projects, a lot of people with conflicting timelines and priorities, and I am the sole hands-on analyst for all of them, working with a faculty supervisor. It's a great group with a LOT going on.



This diagram represents the life cycle of a typical project of ours: We start with a clinical question and turn it into a statistical research question; we develop our experiment design and our analysis plan; I work with our clinical support staff to collect and clean our prospective data and do all the data management. Usually it's just me doing the descriptive statistics, inference, and sometimes prediction based on that analysis plan, and then I work with our clinical PIs to interpret, message, and disseminate those results, usually in an academic manuscript.

There's a huge team involved all throughout the process, and our team is fantastic. But you'll notice, of course, that the stats team is the only one involved at every step.



I think this is true for a corporate data scientist as well, for example - obviously this is a generalization, but again we're working with business stakeholders at some points, domain experts or infrastructure teams at other points, but we're the ones involved at every step.

What happens here is that since we're the ones who are *involved* at every step, we're also the ones who *manage* these steps - we know the investigators' timelines, for example, and we've talked to the domain experts, so we know better than the stakeholders do what a missing data point might mean. We're also the ones juggling multiple priorities, so it's partially up to us to figure out what project should take priority when. In short...

when there is no project manager...

We
become a project manager

bit.ly/jlt-wsds2018 • @jent103

...if there is no formal project manager, often WE become a project manager.

Sometimes there are others who shoulder part of this responsibility; for example, in academia, often the clinical investigator is in theory the project manager of his or her grant. But what happens in reality is that that PI doesn't work on the daily data collection of a prospective study, probably because he or she is busy trying to get the next grant or taking care of patients, and thus the statistician knows more than the PI about the intricacies of the data because we're the ones talking to the research coordinators. Or, we know better than the research coordinators what the overall aims of the study might be, because we're the ones who developed the statistical analysis plan. Being so involved in every aspect of this means we often end up managing it by default.



THIS HAS ADVANTAGES!

- Deeper project/domain knowledge
- Stronger relationships with teammates
- Better understanding of raw data
- These lead to stronger analysis, interpretation, communication
- Many of us enjoy it!

...AND DRAWBACKS



- Project management is a **lot** of work...
- ...which is not typically considered for career advancement
- Doing this well may actually hold us back, if it takes time from “promote-able” tasks & skills development

Link: Tanya Reilly's excellent talk on "glue work"

bit.ly/jlt-wsds2018 • @jent103

This has a lot of advantages! We get a deeper knowledge of the data, the project, and the domain; we develop stronger relationships with, say, the infrastructure team, which makes things more efficient; all these things lead to better analysis, better interpretation of results, and more effective communication of those results and the clinical or business impact. Plus, a lot of us genuinely enjoy it! - I get a great sense of ownership and improved efficiency when I'm this involved in a project.

On the other hand, though, project management is a lot of work, as any project manager will tell you. And in our field, that work is typically NOT considered when it comes time for career advancement. In that case, doing project management well might actually hold us back professionally, if it takes time away from those “promote-able” tasks and professional skills development. Tanya Reilly is a software engineer who has a fantastic talk on doing this “glue work”; a lot of it is applicable to us as data scientists as well.

strategies for success

that have worked in my context 🙌

YMMV:

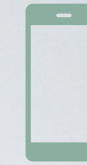
take what works, tweak what doesn't

bit.ly/jlt-wsds2018 • @jent103

So to take advantage of doing these tasks, and also to keep them in proper perspective, we're going to talk about some practical strategies for success. These are the specifics that have worked in my context; your mileage may vary, so I hope that you'll be able to take these ideas and tweak them to work in your specific context.



COMMUNICATION



Regular team meetings

- Group discussion of goals, projects, timelines means everyone understands rationale
- Facilitates updates on roadblocks, successes
- Allows all team members to learn from each other

Public list of priorities

- Trello board, Slack channel...
- Gives stakeholders a clear picture of when their projects will be addressed
- Facilitates conversation about shifting priorities as needed

bit.ly/jlt-wsds2018 • @jent103

The first thing, always, is communication! These two strategies have really improved our group's communication. When I first started with the group, I would get individual requests from PIs with no coordination between them; as a new graduate, obviously this was kind of a nightmare scenario. Several years ago, we started a weekly meeting of the biostats team and all the principal investigators - all the primary project stakeholders. This is incredibly helpful for getting our team all on the same page, for letting people know why I can't work on their project until X date, and for coordinating our priorities and timelines. It also allows every team member to learn from one another, which is particularly important in an academic environment like ours - the research fellows are pretty quiet in this meeting, but they learn a lot from seeing how the clinical faculty think about grants and work with us, for example.

We also keep a public list of priorities available so that all the PIs know what projects are on my list, and have a sense about when I might get to them. This also helps facilitate conversations about when timelines might need to change. We use a Slack channel for this that I update after every week's meeting; you might have a Trello board or something else, but I've found this public list to be really helpful.

FOCUS

bit.ly/jlt-wsds2018 • @jent103

The next few strategies we'll discuss all revolve around focus - keeping projects aligned with our mission and preventing them from sneakily getting out of hand.

KEEP THE MISSION IN MIND

What are your team's goals?



What projects/analyses will move you toward them?



Limit your to-do list to those projects!



bit.ly/jlt-wsds2018 • @jent103

The most important thing, I think, is to keep your mission in mind. What are your group's goals? Revenue, grants, patient care - you can have several, of course, but delineating them gives you criteria for which projects to take on and which should take a backseat. Our group is full of ideas, which is fantastic, but we have definitely had times when we all thought "this idea is great," and I worked on it for a month or two, and then it sat around for months or years because no one had really taken ownership of it - it didn't truly move us toward our goals. So the first step is to choose your projects carefully.

DEALING WITH (SCOPE) CREEPS

Problem:

Scope creep (*n, my version*):

Project gets bigger and bigger without explicit intent 🐱

One solution:

Use your project/statistical analysis plan as a tool 🔧

- Record original intent
- Document & version every change



- Additions/changes mean updating the version; this might be okay, but process forces you to think

bit.ly/jlt-wsds2018 • @jent103

Once you've determined a project is worth taking on, and you've defined the parameters, we all have to worry about scope creep. This is what happens when someone says "oh, as long as you're in that database, can you look at this too?" Or "Oh that is awesome... wait, what happens if we do it THIS way?!" Again, great ideas, and sometimes worth incorporating, but sometimes they just derail us.

We've successfully used our statistical analysis plan to handle this. We're about to have a major clinical trial published, and particularly in that scenario, the SAP is a big deal - we registered ours on the Open Science Framework before we broke the treatment blind, so that it's very clear that we didn't do any manipulation of our outcomes after we knew the treatment assignments. This had a very practical purpose too: whenever someone got a clever idea for more to do with the data, or a different way to analyze things, we were able to say "that's a great idea. We'd have to update the SAP, and it would be after we broke the blind. Is it important enough for that?" And so far the answer has always been no. This means, in our case, that we were able to finish the primary outcomes without being distracted by these side questions. If they're important enough, they'll pop back up, we'll write another plan to deal with those, and revisit the process.

Generally, planning, and then documenting the changes to your plan, is incredibly helpful for setting your intention, for making sure new folks understand the goals and methods of the project, and for putting a process in place that helps you stick to your overall goals.

FOCUS vs AD HOC PROJECTS



MONDAY

"I'm going to work on this big clinical trial! It's going to be awesome!"



FRIDAY

Spent my week on 18 revisions, bug fixes, "can you just tell me how many..."

Results:

1. Group is only inching toward goals, rather than making major progress
2. I am slowly being driven mad

bit.ly/jlt-wsds2018 • @jent103

Now, having the best plan in the world doesn't help if you can't actually execute it. I don't know how many of you can identify with this, but there are many weeks when I come into work thinking I know exactly what I'll be working on, and end the week wondering just exactly where it went. Getting "distracted" - and I'm using air quotes on purpose - by these smaller requests can be kind of maddening, and even worse, it keeps the group from focusing on the major projects that are going to move us toward those goals we set.

FOCUS vs AD HOC PROJECTS



Goal:

Make progress on major work that moves us toward primary mission, while remaining responsive to also-important ad hoc requests

Our specifics:

- **Two** weeks working on major project (eg, primary outcomes of clinical trial)
- **One** week on smaller needs (reviewer responses, bug fixes...)
- Repeat! ↻

Results:

- Major projects done more efficiently
- Smaller projects done in reasonable time frame, communicated well
- Everyone's needs handled
- I feel more satisfied with tasks & output

bit.ly/jlt-wsds2018 • @jent103

We've instituted a system inspired by some software development strategies: I spend two weeks of my time working on the "major" projects that require lots of focus and move us toward our group goals, and the other week on the smaller projects that are still important, but can be stacked and worked on in batches. These might look like those "hey, can you tell me how many people..." requests, or journal revisions, or changing color schemes - things that do matter, but if we give them immediate attention every time, might actually distract us from our mission.

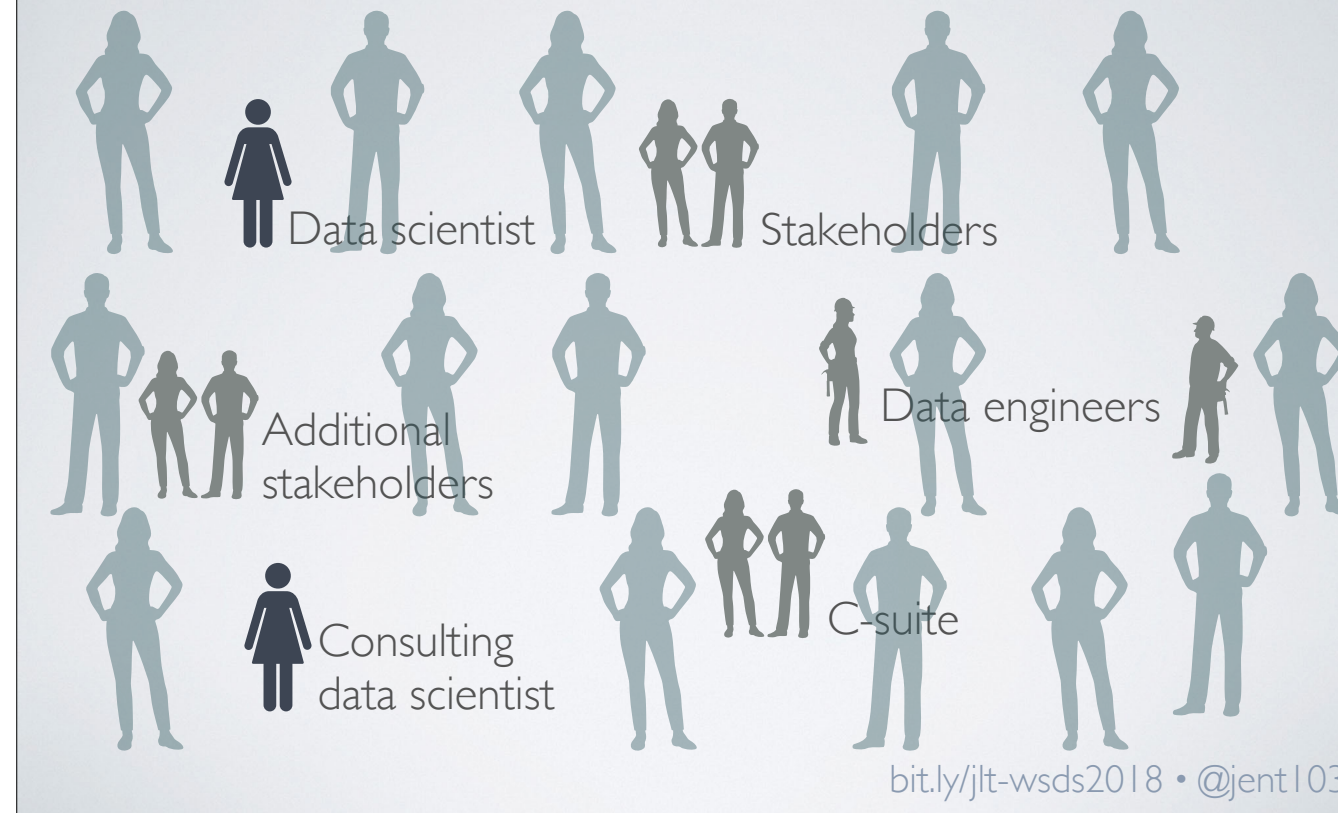
We've been using this system for probably six months now, and I am happy to report that it's made me much happier, because I'm not multitasking on eighteen things at once all the time, and everyone else seems to be at best enthusiastic and at worst neutral about it. It has allowed us to push a major clinical trial through to publication in a reasonable time frame, while still allowing the daily business of clinical research to take place. In short, I'm a big fan of this system. It may not work just like this for everyone, but the concept of delineating some time as "focus time" and some as "ad hoc time" is one I think a lot of us will benefit from.

PROCESSES

bit.ly/jlt-wsds2018 • @jent103

Our next couple of strategies will focus on processes - streamlining a group's work so that details don't get lost and chaos doesn't ensue.

WHO MAKES DECISIONS?



In your typical project, who is involved in making decisions about analysis plans, business questions, communication strategies? Of course hopefully we, the data people, are; there are probably stakeholders of some kind involved too. There might be some additional stakeholders - if my primary collaborator is a research fellow, there's probably a faculty mentor involved, for example. Maybe we have a database team or other infrastructure folks who help decide about data storage. Maybe we have a statistical colleague with expertise in a particular area... maybe we have a C-suite or a department chair we need to make happy.

Eventually, there are so many people in the metaphorical room, and if we're not careful, every decision needs to be run by all of them. This creates chaos, not consensus.

WHO IS...

Task: Developing an SAP

Responsible



Biostatistician



Research fellow

Accountable



Biostatistician

Consulted



Faculty mentor

Informed



Additional faculty, research coordinator, DB team

more:
bit.ly/dpm-raci

bit.ly/jlt-wsds2018 • @jent103

Enter the RACI matrix. This is a project management tool that we've found pretty helpful recently in terms of streamlining decision-making processes. For any given task, you define who is **responsible** for executing the task; who is ultimately **accountable** for the task; who is **consulted** on decisions related to the task; and who is simply **informed** about those decisions.

To demonstrate how this might work, I've outlined the task of developing a statistical analysis plan. Obviously I, the biostatistician, am partially responsible for this statistical task; my hypothetical research fellow is also involved, since there are clinical decisions that need to be made. But ultimately, this is my zone, and I'm accountable for delivering the final product.

My research fellow probably has a faculty mentor who will also be consulted when we're making those clinical decisions or refining the research questions. And we might have additional clinical or statistical faculty, maybe a research coordinator, maybe some infrastructure folks who will need to be informed about the decisions. But that's it! The entire world doesn't need to be consulted before we move forward - this is the core team responsible for the product. Once we know this, we know who we need to wait on for decisions and who we can move forward without on this particular task.

DOCUMENTATION

Project request form
for gathering initial info

- Goals for analysis
- **A**ccountable &
consulted team
members
- Logistics



Potential roadblocks discussed *before* data work gets underway;
clearer vision of decision-makers, end product, & value added

CIBS Center Biostatistics & Data Management Core Project Request

We are excited to collaborate with you! Before submitting your request, please read through the Biostatistics Core Onboarding Guide (NOTE: currently in progress; will be linked here when ready). This document answers common questions and explains how best to work with the biostats core to efficiently produce the best research possible.

These processes are in place to help us manage an ever-growing group of investigators and their needs. The information below will help us work with you to understand, prioritize, and execute your project. We expect these details to evolve to some extent during the course of the project, but we've learned from experience that thinking through them before starting makes the process much smoother.

In progress:
Onboarding
docs for
new team
members

Deadlines 🕒 What's already available? 📄
Funding source 💰 Target product 🎯
IRB permissions 🔑 Data source(s) 💻

bit.ly/jlt-wsds2018 • @jent103

We incorporate that RACI matrix into the information we gather with our project request form. We've started doing this recently for new statistical projects; it helps us learn the reason for the project, whether that's pilot data for a grant, or a new manuscript; it tells us who is **accountable** for the project and the additional team members who will help steer it by making decisions. It also collects several pieces of logistical information. In our context, that looks like deadlines; any manuscript drafts or table shells that are already written, and what the end product will look like; in our work, thinking about funding source is necessary - where is the money for this analysis coming from?

We hope that asking about these potential logistical hurdles beforehand will streamline the process - for example, if I'm not listed as key study personnel with the institutional review board, I can't legally access the data, which is obviously an issue. It's better for the investigator to think about that beforehand than when I'm ready to start working on it.

In an academic environment, we've also found that asking PIs to write a draft manuscript or abstract shell before analysis is helpful in several ways - we get a better picture of the research question, and once the analysis is finished, there's less time before the manuscript actually gets written.

We've also started working on some onboarding documents for new team members, with information like how and when to contact the biostats team, where our documents and data are stored, what we expect both of our investigators and of ourselves. Coming at this with a "we're on the same team" approach is really helpful, and we're hopeful that collecting all this information in one spot will be useful for our new clinical and statistical team members.



SYSTEMIC CONSIDERATIONS

bit.ly/jlt-wsds2018 • @jent103

We've talked about some very practical strategies for getting things done efficiently, but there are also systemic considerations. I think all of us can agree that some type of project management is important, but how does that fit within the typical organizational structure in the absence of that official role?

- Can we do a better job at **recognizing** these professional skills?

Relevant: Manisha Desai's WSDS2017 talk on recognizing team scientists in academia

- Varied career ladders - industry currently does this better than academia IMO
- Update typical promotion criteria
- Do *you* have ideas?
- Offset time spent on PM tasks by asking for **protected time** to either perform those tasks, or to focus on promote-able tasks
professional development, developing group infrastructure...
- Are you a **manager/supervisor**? Do you see this happening? Step in!
- **Know what is valued** by your organization, your career path - and *yourself*



bit.ly/jlt-wsds2018 • @jent103

First, I wonder if we can do a better job *recognizing* these important professional skills. A typical career ladder, especially in academia, doesn't officially recognize these collaborative professional skills - this is something Manisha Desai brought up last year in the context of rewarding team scientists. I think industry does a better job at this, at least in theory; for example, I've seen titles like "statistical systems manager." I'd be curious to hear from those of you in industry about your take on it, and from anyone about ideas to officially recognize folks who have worked hard on these skills.

One idea that may be realistic for some and pie-in-the-sky for others is to ask for protected time - that might be protected time to actually do these project management tasks, in other words your boss recognizes that at least 10% of your work week is these logistical tasks. Or it might be asking for that time to be protected for professional skills development or developing group infrastructure - specific tasks that will both benefit your team and lead to more promote-able skills for you. Not every situation will be excited about this request, but if you think it would be a possibility, I think it's worth a shot - even having the conversation might bring awareness to this "glue work" you're doing.

If you're a manager or a supervisor, and you see folks doing a lot of this glue work without recognition or compensation, please please step in! Especially here at the *women* in statistics & data science conference, a lot of us aim to please by default, and a lot of us want to fix inefficiencies and make things the best they can be. That is great, up to a point, but if we're doing tasks that aren't actually valued in any way other than maybe a verbal thank you, it will hurt us long-term. A lot of us don't realize that when we're fresh out of school and wanting to make a good impression - I know it took me a long time to realize how much of my time was spent on things that weren't really considered by my promotions committee.

And finally, I think it's so important to understand what is valued by your organization and the career path that you're on - and to know what you yourself value. Ideally, if you enjoy these kinds of project management tasks and work on developing these skills, you'll be at a place where that type of professional development is valued and recognized. But if you're in an academic tenure track position, most likely no one is asking you how efficiently your collaborators' first author papers were written. That's a

ACKNOWLEDGEMENTS & RESOURCES

- [VUMC Strategy & Innovation Office](#) for project management training
- Jesse Mostipak & Sharla Gelfand for nonprofit & industry perspectives 🐜🐜

Related reading!



[Trey Causey on the data scientist as product manager](#)
(product != project, but many similar concepts)

[Tanya Reilly on “glue work”](#) (h/t to Sharla!)

Roger Peng on [balancing resources, analysts managing the flow of information](#)

[bit.ly/jlt-wsds2018](#) • [@jent103](#) • [jenthompson.me](#)

Finally, I want to thank some folks! As I mentioned at the beginning, I work at Vanderbilt, and our strategy and innovation office has some great training workshops including some on project management, which is where I learned some of these techniques. If you're at a large organization, I'd highly recommend looking into options like that. Also, I would love to thank Jesse Mostipak and Sharla Gelfand; I've been in academics my entire career thus far, and they helped make sure this talk would generally apply to all kinds of data folks.

I've included some further resources here, along with the link to the slides and my web site if you'd like to talk further - I would love to hear your questions and ideas. Thank you!