

The way we're working isn't working

Why we need to improve the way we work
- partly based on true stories

Request/invitation to collaborate

Dear Hans Olsen,

I have some data that I would like to have analyzed..

Ask for data

Dear Grete Jensen,

That's great! Would you send me your data?

Try to give data

Dear Hans,

I have now sent an email with the data

Dear Grete,

I didn't receive anything..

Dear Hans,

It seems the email was too big..

Dear Grete,

*would you be able to provide ftp or http access
to your data?*

Dear Hans,

ftp? http?

Dear Grete,

Let's instead try www.filemail.com

Dear Hans,

Hurray - I got it to work. Thanks a lot!

You will get a separate download mail..

Starting development

Dear Grete,

I have now successfully downloaded your data. I will start developing and let you know when I have some results.

Incorrect data

Dear Grete,

Sorry for the long delay. I now got the functionality ready, but unfortunately it turned out that all the data were not valid - some regions went beyond chromosome size..

Dear Hans,

That's okay. I'm sorry for the mixup. I am on a conference now, and don't have access to my data here. I will look into the problem when I'm back.

Dear Hans,

I am now back from conference, and have resolved the issue with incorrect data.

I have now sent you a new version.. Is that okay?

Ready for analysis

Dear Grete,

the data was now working well and I was able to do the analysis. Please find the results attached.

Need further details

Dear Hans,

Thanks for the results. I was wondering about some values of the green curve, though. Could you send over some underlying numbers for these results?

Dear Grete,

Sorry for the long delay. I didn't find the underlying details, and had to re-run the analysis. Please find attached the numbers you wanted.

Modify analysis

Dear Hans,

No problem. Thanks for the details. The numbers were indeed unfortunate. I think we will need to make some modifications to the analysis. What did you assume/select...?

Dear Grete,

The parameters of the analysis are as follows:

...

Dear Hans,

Could you try value 40 of parameter B?

Dear Grete,

here are results with value 40 for B. Is that good?

Dear Hans,

that's better:)

*Could you also try runs with values 4,5,6 and 7
for parameter D?*

Dear Grete,

*Attached are the results with values 4,5,6 and 7
for parameter D. Do they look good?*

Dear Hans,

Sorry to bother you again - but which of the results were for which value of the parameter?

Dear Grete,

Sorry, I forgot to specify. They are as follows: ...

Finally, we're there!

Dear Hans,

It now looks good:)

After a dozen of further iterations on email we should be ready to publish!

Dear Hans,

The article is now submitted!

**But,
there's also a revision**

Dear Hans,

This is Grete again. Long time no see.

We just got the article back from review. They liked it, but request that we run a similar analysis on some other data set (attached).

Would that be okay?

Dear Grete,

That is great!

*I'll do a corresponding analysis on the new data
and come back to you!*

Dear Grete,

Results are on the way. I didn't remember exactly how to use some of the previous functionality, but I'm looking in to it now, and will get back to you..

Dear Grete,

Unfortunately, I haven't been able to track down exactly how we analyzed the previous data. I found some scripts lying around, but don't know which I used for the final analysis, nor exactly how I used them.

It was now able to make it run, but I don't get the exact same results on the previous data..

Dear Hans,

That was a bummer.

I guess we should just use some settings that give as similar results to the previous ones as possible, and hope the reviewers doesn't notice.

The revision deadline is anyway approaching..

[A dozen emails later]

Dear Hans,

The article is now re-submitted!

Dear Hans,

The article was accepted!

Time for champagne:)

And then someone may
even read the article..

Dear Hans Olsen,

*This is Grete Jensen. Do you remember me and
the article we published around a year ago?*

*I got contacted by a group that was interested in
our method. Would you be able to send over the
analysis pipeline you used?*

But that's anyway the end of it

Dear Grete Jensen,

Yes, I remember you.

Unfortunately I do not remember much of the method. I have some scripts lying, but they don't seem to run anymore, as they required some programs I had installed on a previous laptop.. And to be honest I don't remember how to use them. I think we should just wait a bit and see if they just forget about it. If not, we can handle that then..

An alternative way to work

How some structure,
along with supportive functionality,
can make the process faster and smoother

Request/invitation to collaborate

Dear Hans Olsen,

I have some data that I would like to have analyzed..

Ask for data

Dear Grete Jensen,

That's great! Could you please upload the data to a Galaxy history and share with me? And please validate your data through a simple analysis like counting of data. For more details see our help page with all necessary links: [www...](#)

Dear Hans,

I have shared with you a Galaxy history with all data at: www...

(by the way I discovered some data issues, but it all seems to be working correctly now)

Dear Grete,

I have now developed a tool for the analysis you were interested in, available as tool “X” at [www...](#)

I did an example analysis (shared with [URL:...](#)), but you probably know better how to select appropriate parameters (selectable in the tool)

Dear Hans,

Thanks - the tool worked great! After some exploration I was able to find suited settings.

I have attached a manuscript draft - do the Methods description (based on your tool help) and the results seem appropriate?

Dear Grete,

The manuscript looks great! I have made some suggested changes to the attached version.

Dear Hans,

*The manuscript is now submitted.
Fingers crossed:)*

Dear Hans,

This is just to let you know that the manuscript came back from review with quite positive remarks. I just need to do a similar analysis on some new data, and will then resubmit. I don't think I will need any further input from you at this stage, and will let you know otherwise.

Dear Hans,

The article was accepted!

Time for champagne:)

Dear Hans,

This is just to drop you a line to notice you on the just published article that cited us. They used our analysis methodology on some new data, and got exciting results. The group behind the publication is a very interesting one, and obviously they liked your methodology:)

Surely you're cheating,
Mr. Feynman!

And what's really the difference anyway?

The difference

- **Sharing is done using a URL**
 - The data is the same place throughout the process
- **Data is validated straight from the start**
 - Any faulty data are discovered early and by the biologist
- **Exploration of settings is done by the biologist**
 - The biologist tries parameter settings and inspect results
- **The analysis pipeline is automatically tracked**
 - All steps, programs, parameters and results are tracked

The alternative way in practice

How the described process can be achieved using
Galaxy and some simple protocols

The concrete instantiations

- Shared data:
History: “imported: DNase_for_psychiatric”
- Analysis:
History: “MS-DNase-results”
- Public access:
Page: “MS-DNase-analysis”