DCC implementatienetwerk - How to organise a DCC and support

Tuesday 29 June & Thursday 1 July

As a DCC we have the task to organise research (data) support and help researchers to perform high-quality research by providing them the necessary tools for good data management, FAIR data practice and research reproducibility. Tools consist of infrastructure, policies and guidelines, support services, training and communities. We all have similar challenges and we are all working on solutions to solve the challenges. We are also aware that infrastructure is not enough (although good infrastructure does help!). We need engagement from researchers and we need to connect researchers and supporters.

RUG, UMCG and VU presented briefly their approach in organising a DCC and support. Three different organisations with the same challenges and, slightly, different approaches. In the presentations three questions were raised to be discussed in breakout groups. These questions reflect the main topic of the meeting: how to organise a DCC and support? Each presenter provided a central question which was discussed by a group.

1. How to address researchers?
2. How to organise support?
3. How to find a balance?

The idea of the discussions in the breakout groups was to collect and connect. Collect brilliant ideas, approaches, tips & tricks. And connect these with the participants' challenges. Below you will find what we have harvested in 2x3 breakout sessions. This is given in bullet points as it is undoable to reflect the full discussions in the groups. Overall we can conclude that these were extremely valuable short discussions, resulting in some great ideas and insights. But also the conclusion that we are not there yet and that we still have challenges which we should address collaboratively.

Break out Group 1: How to reach researchers?

In this group we focussed on the intended audience, the researchers. Who are they (all researchers, only PhD's?). What is the size of our intended audience (do we all have similar numbers?). How hard should we market research support, if at all? And how do we explain the apparent mismatch between the volume of research support delivered vs. the size.

Tuesday 29 June:

- Ensure that the research support 'problem' is carried by the whole university. Make it a bigger issue than just of the DCC so that more people can point towards the 'solution' (first build the network - then tackle the visibility issue as a network).
- Look at the kind of questions, more complex questions mean more time to answer them and a different role (advise instead of support). Numbers are not everything (quality over quantity).
- Decanen en wetenschappelijk directeuren zijn een belangrijke doelgroep
- Veel cursussen zijn vrijblijvend.. dus een hele andere insteek ook daarvoor

Thursday 1 July:

- Samenwerking HPC en DCC + lange termijn investering in Data Stewardship (ziet exponentiële groei in aantal vragen - TU Delft)
- Vroeg in de onderzoekscarrière aanwezig zijn (onboarding PhD's)
- Aanbod niet vrijblijvend (bijv. studiepunten toekennen) - incentives to participate
- Aanbod concreet maken (meer dan informatievoorziening, bijv. workshop met RDMP als deliverable)

Break out Group 2: Central vs embedded stewardship

In the 2nd group we tried to find an answer to the question of positioning data stewards. How are institutes co-ordinating the central and embedded data stewards? Is there an ideal balance? Are data steward tasks shared between other research-support function groups? (e.g. research coordinators, data managers, data scientists). How can we make sure central and embedded are on the same page?Are embedded data stewards the solution or will they in fact contribute to "hidden/invisible" research support?

Your ideas, approaches, tips and tricks
Tuesday 29 June:

- Large variation in 'data steward' capacity. But also different definitions
- Try to build connections to 'unknown' data management supporters in research groups
- Bat in the henhouse: don't focus on anonymous IT solutions for supporting researchers
- Personal approach/contact is vital? Many researchers do not like to communicate through 'anonymous' platforms/general emails (tip: always discuss with the researchers the need for a complete dossier in the CMS, which can be read by all DSs. Suggest always putting in cc: the general email address -with call number in the title- as additional to emailing individual DSs, if direct contact is desired)

Thursday 1 July:

- How do we reach the embedded data stewards?
- Do we know what is central and what is embedded? Is it the case that central is generic and embedded is discipline specific?
- Does the embedded data steward who actually touches the data exist?
- Embedded data steward for 0,1 fte doesn't even have the capacity to work with metadata and data
- Initiatives like ZonMW are not matched with the capacity
- Are there opportunities for researchers to become embedded data stewards and thus stay in research but avoid competitive career track towards becoming a professor
- Wageningen: the tactic was finding out researchers who are already doing RDM part-time. This created 70 data stewards - it was agreed with the board that these people will get part of their time reimbursed. Interestingly, the data stewards are now trying to broaden their tasks (by adding more research groups for examples) and get more FTEs to do RDM more seriously. The DCC hired a coordinator to bring together the data stewards and to coordinate them. The big problem is continuity - researchers are often on temporary contract and need to move on - which is not what you want for a data steward

Break out Group 3: How to find a balance?

In the 3rd group we talked about balances. Actually we addressed two different questions here. One was the question on how to find a reasonable match between researcher's needs and organisational requirements and regulations? How to provide the support that is actually useful to researchers within the organisational policies. The second question was related to the quality of support. How to find the right balance between providing researchers with the best possible answer to their questions and providing this with a reasonable time knowing that there is a limited staff. Related to this is the question whether we should collect metrics about support and measure quality of support.

On Tuesday we addressed the first question and harvested the following:

- We should involve researchers from the beginning. Ask them to prioritize what is important to them. This helps in giving the best support, tailored to their priorities. It also delivers "good will" for rules and regulations they simply have to follow
- campaigning is important. If you have (new) facilities, promote them in all possible ways. The example of a poster about data storage was given
- it is important to reduce the distance between researchers and support. embedding data stewards in faculties and ask them to come up with a plan and common topics helps a lot
- organise thematic meetings based on researchers' questions and organised together with researchers. This helps in making them a bit of the owner of the topic/question
- These topics should be very concrete like how to use cloud computing, what type of storage should I use, where to archive data, what about software archiving.
- make an inventory of questions within faculties
- in a community you can appoint experts in different domains

On Thursday we addressed the other question:

- the balancing act of matching quality with response time is especially an issue when we have to service a large community. When having only few scientists there is no real problem.
- Otherwise we struggle with the problem of how to match questions and answer with each other and how do we organise this
- We agreed that we need a "knowledge base" in which we can collect information from services like TopDesk and communication channels like Slack
- But we also agreed that development is difficult as we find ourselves in the position that institutes have many communication channels and it is not always clear which channel to use. Information about RDM is often fragmented
- TopDesk provides an option to publish answers from its internal knowledge base. We should explore these possibilities together
- about metrics it was stated that we not only have to measure on output (how many answers did we provide, how many trainings did we give), but also on input: how many faculties do have an up-to-date RDM-policy, how many researchers follow a training course, how fte are invested in data support per faculty, ...
- Then we also will be more able to get a hold on what types of questions coming in, what are the answers. And we could measure what types of questions come in (what topics, complexity). Based on this we are able to act and organise f.i. (thematic training, webinars, etc

- Finally we very much agreed that providing good and structured information is key to success.