For consideration at DCP 350 meeting on 20 January 2020.

From Rob Nickerson, with comments in red from Lisa Waters

I've proposed a few next steps some of which I am happy to lead on, others require working group input. I'm not 100% sure what this means for our timeline but given that we have already slipped from February 2020 to April 2020 this data won't land in time for the next FES/EMR cycle anyway. I'd rather us get the details right – whilst progressing the legal side – so that we end up with the right dataset for the main use case in the end. This could be a good test case to prove that the industry can follow the good data guidelines being developed by Energy Data Taskforce / Energy Systems Catapult (Energy Data energydata@es.catapult.org.uk).

Co-location

My main concern is about co-located sites as these are becoming increasingly popular. From a security of supply point of view it is important that we properly understand co-location as we will not be able to assume full output from each component if the total site's export capacity is limited. One option seems to be to list the sites primary type (e.g. "wind" for a wind + solar + storage site). However this misses a lot of useful information about how the site can generate and store power. It is for this reason that I proposed a 2 table solution in which we record generation asset details in a separate (but associated) table. I've reattached my file as I am aware than the worksheet tabs may have been hidden in the original.

Next steps:

- It would be good to get some input from the DNOs to help find a way forward here.
- In particular, how would some of the co-located sites being connected now look within the currently proposed register.
- We may need to ask Scottish Power Distribution as I'm aware that they have a lot of collocated sites in the pipeline.

LW – in the connection agreement you have to say the type of kit on the site, so can we not have something like this:
Site Name X – Technology Type A + B + C – connection MW?
Ideally we would want the capacity of each type of kit – is this in the connection agreement?

Changes in capacity

A further concern I have is around the capacity columns, in particular the use of the "Capacity Increase / Decrease (MW)" column. I've found this to be problematic in the TEC Register and explained earlier in the life of DCP 350. Aware that not everyone saw my explanation so sharing it here:

> <u>Version control (MW increase/decrease)</u>
> As a user of the registers, I am not a fan of the triplet of columns related to the project size ("MW Connected" / "MW Increase/Decrease" / "MW Total"). This can be very confusing and would not be best practice for a modern database design. I've already seen it inconsistently used in the TEC Register thus requiring manual work before the data can be safely used. If some manual interpretation is required it is just about possible with a register of 500 sites (with TEC), it is simply not going to be possible when we have 3500 sites (Distributed Generation). In my opinion it is much easier to have the MW entry with a Valid from date and Valid to date. The register would then hold multiple rows for an individual site which has

increased or decreased its capacity. As a data user, I can then filter for those date periods I am interested in. (The Valid to date could be automatically populated as in my template file included within the original change proposal).

I think we need to be really careful with what we decide here. As noted it just will not be possible to manually review 3500 sites if the data is input inconsistently. My preference is to use new rows to record the increase/decrease in capacity (each row having a start and end date). I would be interested to hear your thinking here.
LW – I agree it is tricky, but you get use to it on the TEC register! I think the increase or decrease may well be dates that keep getting moved, so I don't see how NG could really use more than the actual connection size.

Next steps:

- Other members to provide their view on what makes most sense.
- Option to test this thinking with an expert in the field of data / data table design (e.g. can we lean on the data work being done by the Energy Data Taskforce / Systems Catapult here?) (Energy Data energydata@es.catapult.org.uk).

"Technology / Plant type" and "Resource"

It is definitely the right thing to do asking for the detail as this is increasingly important. I am aware however that the DNOs may not be able to report it straight away. I think the compromise position here is that we ask for this to be completed to the best of their knowledge. Over time the quality would be expected to improve.

Next steps:

- Are DNOs/IDNOs happy with this approach of best endeavours at the start with quality improving over time?
- If yes, can we look at what list we think makes most sense in the long run. Rather than jump in at this stage I think it would be good to first understand how you (Dylan and Richard) came to the list that you have proposed. Can you share some details please?

LW – agreed – can we see the list?

Grid Supply Points

For Grid Supply Points can we please use an agreed set (i.e. a drop down list). For this I propose the Elexon list (as included within the reference network mapping statement). The reason I ask is because I have learned the hard way that GSP names (and even which substations we treat as GSPs) seems to differ even within one organisation.

Next steps:

- Are the DNOs happy with the proposal of using a single definitive list of GSPs? And that this list is best aligned with the settlement process (Elexon)?
- If so, I am happy engaging with Elexon and DCUSA on this.

LW drop down seems sensible to sort by, but need a map of the GSPs as well.

Site address

Please consider adding site address as part of the locational details. This is useful to allow the data to be associated with other datasets. It also allows wider use by parties who do not have the knowledge of the electricity network (e.g. may not know the location of GSPs).
LW – I thought we had address. I prefer it to the OS reference as that seems to not link site in some areas, but is used in the CM.

Response to above from Rob Nickerson

On co-location it is good to know that in the connection agreement you have to say the type of kit on the site. I agree to that we would want the capacity of each type of kit in our register. The challenge is that the register must also be easy to use and interpretable using a software script (with over 3500 sites we won't be able to look at each entry manually if the data is recorded in an inconsistent way). I welcome the thoughts of this group as to how to solve that problem. When I thought about it I came to the conclusion that you need two tables – one with the site summary details and a second with the details of the kit (one piece of kit per row). I'm re-attaching that again but if other solutions are possible please let me know.

Seeing how dates move is actually an interesting observation in itself and is something we could analyse to understand typical delays. We recently had RTE (the French TSO) present to us on some analysis they did based on how the nuclear sites typically re-opened after maintenance an average of 4 weeks later than their original plans. RTE used this knowledge within their security of supply analysis. Just an example to show that all data should be considered even if at first value it doesn't appear useful.

By the way my team have reported data inconsistencies in both the TEC and Interconnector registers over the last few weeks. These have now been corrected. I make these comments, not because I want to criticise my ESO colleagues who work on these registers, but because I feel it is important to learn lessons and make sure we don't repeat them here. As noted we will not be manually looking at 3500+ rows of data so we need to ensure this is done right. As it currently stands I feel there is too much risk of inconsistent data entry.

>but need a map of the GSPs as well.

What do you mean by "map"? I ask because the ESO publishes latitude and longitude coordinates for each GSP – is that sufficient? Or are you meaning a mapping between DNO, GSP, and lower voltage sunstations?

Comments regarding consultation from Rob Nickerson

Further to my email below on the consultation, I have now read the legal text. A few small comments:

- Do we need to specify that the data is machine readable – e.g Excel format?
- Given that we will be downloading and combining these using a computer script it would be nice to include reference to the fact that the DNOs/IDNOs should publish the register in a non changing location (permanent link/"permalink") – e.g. www.example.com/embbedded-register.xlsx. There may need to be some wiggle room for occasional changes e.g. in the case of website refreshes/branding changes.

- In light of the Energy Data Taskforce recommendations, do we want to specify that the data should be accompanied by a licence setting out how the data can be used and that this licence should be an open data licence? See for example the Open Government Licence.

Regarding Lisa's email comments (18/12/2019):

> Do we add a list of data to be in the ECR? I think so as we then what the DNOs are meant to be doing;
> Where are the definitions? In a sheet on the ECR? I think it makes sense, but could define each field in the legal text?

My understanding is that we are not putting the detail in the legal text. Instead we put it in the ECR and then have a process where changes can be submitted to the Panel. If that works then I am ok with that approach.

In regards to the supporting document you provided, do we need to be adding a reference to the ECR is section 34.2.3/34.2.4? At what stage do we get DCUSA/ElectraLink legal advice here? Section 60.12 seems to suggest that we need to get this right otherwise the obligation can by dismissed.