
Un-box the NDR

Myths and reality around National Data Repositories

Setting up a National Data Repository requires understanding a country's goals, deciding upon strategic and financial models, and figuring out the appropriate operational processes.

Going shopping, or "Procuring an NDR," is actually one of the last steps in that sequence and possibly the least important. It is not like buying a car from your favorite car dealer. You would have to go into several shops anyway because, SURPRISE, there is no such a product as a "turn-key" NDR that you can buy off the shelf and enjoy. Like it or not, to build an NDR you would have to open the hood, take a look at the engine and get your hands dirty.

An NDR is not like cognac: the more mature, the better. In the last 20 years, dramatic changes have happened to the way information is managed. This is clear to anyone just looking at the Internet. Most of the ideas from the 90's are either outdated or still in the realm of science fiction. You should indeed study the experience of others but do not stop thinking for yourself.

This leaflet contains a few hints and it answers a couple of questions for those who are considering setting-up an NDR. We hope it can stimulate your vision of what a successful NDR would be like for your country.

Who and what is it for?

The goals of any NDR are clear, although quite usually forgotten:

1. To **preserve** the data generated inside the country by the industry, treating it as a national asset of high importance.
2. To **promote** investments in the country by reducing the risks for both new and existing companies, presenting information in a unified and open way.

Should it make money?

An NDR is a important investment for your country. The returns of this investment is something you should measure in a wide perspective. Does it help the oil companies operating in your country to make better decisions? Does it attract new companies?

Perhaps it is sufficient to simply recover maintenance costs and focus rather on the longer term goals of the investment, like most of us normally do when investing money in banks, bonds and funds.

Should I get everything I need from one single vendor?

"Un-boxing" an NDR is very important. Some long-established players still propose, based on the experience of early NDR initiatives, that an NDR should be procured as an "out of the box" solution from one vendor, where software, operations and even infrastructure are bundled together. In this day and age this is an outdated, sluggish, and expensive way to approach establishing and running an NDR. Business tasks and priorities are to be identified. Tools and actors that are best for the job are to be selected to implement the business tasks, and glued together by open standards in software, IT and operations and most importantly by serious commitment and professional project management from the Government side.

The “box” model is very unhealthy. It covers up the inefficiencies between Operations, Infrastructure and Software. And you get the bill at the end, the settling of which would require significant additional funds. Even worse, some “box” model NDRs are granting some exclusivity on their services to the industry for a selected vendor, allowing him to recover costs. This negatively affects competition and pushes an NDR further into the hands of the vendor. Do you still feel it is YOUR data?

What is “operations”?

Wikipedia manages a vast amount of world knowledge and it is manned by 3 people. Why then you need to employ a squadron of specially trained people to operate your databank?

It is a provocative question but it is worth thinking over. If you have established good procedures and good regulations, outsourced the infrastructure and selected good software, all you need specially trained people to do is control the quality of incoming information. And by quality here we mean the content, rather than the shape or format of the data, because that can and should be verified semi or fully automatically.

About Quality Control and databases

Do you need to put everything into a database to guarantee good data quality? Absolutely not!

Quality Control and NDR databases are too-often linked more closely than they should be, but they are two different concepts. Real quality control is often conceptually replaced by the procedures needed to fit the data into the specific relational model used in the database. It is important to understand that QC and a database are not synonymous.

- The fact that the data was pushed into a database does not mean that its quality is controlled.
- Quality Control of the data does not necessarily mean that the data should end up in a database.

Do not confuse the two. Concentrate on controlling the quality of your data files and making sure all the mandatory metadata is populated and correct.

Box the IT

While we are suggesting to “Un-box” the NDR we would suggest to do the opposite with the IT infrastructure. By now it has become a commodity and can and should be boxed. You need to buy bandwidth, CPU and Disk space. You do not need to buy servers anymore.

And more...

So there is a lot of things you would have to figure out for yourself. Our last advice is to go to NDR conferences, share experiences and get ideas. The best thing you can do is to view, plan, and manage your NDR with an eye for clear business planning, professional management and execution.

“As soon as something seems the most obvious thing in the world, it means that we have abandoned all attempts at understanding it.” (Bertolt Brecht)

