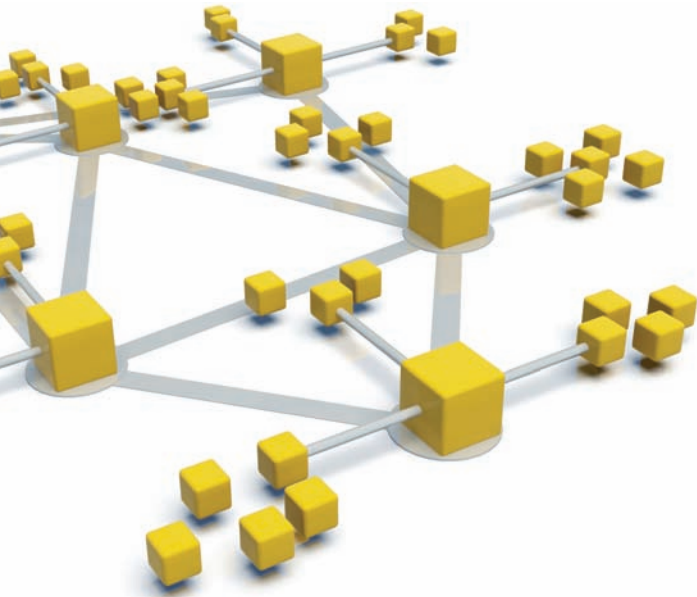


Integrated Data Management in Lab Animal Research



By taking into account a few rules and some simple measures, organisations can get more out of existing software system.

Andreas Staubi

Whether it is used to get better understanding of mechanisms, to reproduce results, to prove a theory, or even to discard a hypothesis, data is a crucial asset in research. Therefore data management is understood to be a very important task in every research program.

In their constant endeavour to make these processes more efficient, research organisations have begun using software systems—and system here can mean anything from a simple spreadsheet to a sophisticated enterprise system—to store and manage the actual research data as well as data which is required for supporting processes and legal requirements.

Sometimes one could get the impression that these software systems are growing within an organisation like weeds.

A variety of spreadsheets, programs, and software solutions can be found for animal ordering, material management, animal stock management, IACUC tasks, breeding, health management, environmental monitoring, and all kinds of scientific applications like LIMS, document management, or nowadays Electronic Lab Notebooks.

The reasons for this colorful software landscape are manifold but often we can hear a similar story: At the time there was an urgent business need to fulfill a particular scientific, regulatory, or process requirement, so somebody spent some evenings and weekends on a nice spreadsheet or database solution or purchased software, and the solution was put in place.

Over time the environment grows without a kind of “meta-plan” controlling it. One software package is bought for purpose A, another one is built using Excel for purpose B, the internal IT solved problem C with another kind of solution and all systems are doing a more or less great job, but do not (and even worse: cannot!) talk to each other.

Historically grown systems were once implemented for very good reasons (even though these reasons are sometimes forgotten by the organisation) and are used by people more or less enthusiastically to keep the business process up and running and to do what is called data management.

AM I GETTING THE SUPPORT I NEED?

Is this application zoo really capable of supporting your data management?

Let's start with a very general definition:

“Data management is the development, execution and supervision of plans, policies, programs and practices that control, protect, deliver and enhance the value of data and information assets¹”.

Can a historically grown system environment support actual data management according to this definition?

Capture and store for sure—that is what all these solutions were built for. Probably they also offer at least some data control and protection, but what about delivery? And what about enhancing value?

Turning the question around: Do the systems in place within your organisation allow users to effectively find the data they put in, to re-use, to re-combine it, to gain useful information and finally knowledge to benefit from?

Sounds too theoretical, more relevant for IT people than for you?

Well, probably it is quite relevant because these questions are directly affecting day-to-day work—your work. For illustration, just a few questions:

- Do you sometimes need to re-enter the same piece of information in multiple systems?
- Do you ever have to spend a lot of time creating reports containing data you know is stored in the diverse systems but you can't get out?
- Do you ever have difficulties mapping data coming from more than one source because it is unclear how the bits and pieces are coming together?
- Have you ever asked yourself what meaning could be behind the particular piece of information you have just entered and what it could be used for?
- Have you ever had difficulties explaining to your team members why something needs to be recorded?
- Have you ever had difficulties sharing data with colleagues within your organisation because of different semantics?

If the answer to any of these questions is “yes”, then, sorry to say, there is room for improvement within your data management system because obviously your software systems neither deliver what you require nor do they sufficiently support you in realising the value.

WHAT DO I DO NOW?

But what can one do with that diagnosis? What are the options?

Of course there is the one-system-does-it-all-approach—one single software platform, maybe coming with specific but fully integrated modules, helping you deal with protocols, animal ordering, animal management, all kinds of test data, doing math and statistics, and finally producing exactly the outcomes and reports you and your stakeholders like to have.

In other domains, e.g. production and financials, this approach seems to work quite well, however, in *in vivo* R&D this idea is followed quite rarely. Of course there are software programs available to manage a large subset of these processes, but the highly specific requirements depending on the area and the type of research and the relatively small size of the market seem to hinder a breakthrough in development and use of such one-size-fits-all systems.

A more common and promising way to address the problems of taking benefit from the data is systems integration². The term simply describes the idea of linking systems, enabling them to talk to each other in order to automatically move, share, and retrieve data between and from different software systems and processes.

Even though the idea (and the need) is as old as the use of software in R&D, it keeps entire IT departments busy.

CAN I WORK WITH WHAT I HAVE?

So, if you are not an IT expert and you don't have the budget to purchase a new software package or integration project, what do you do? Is there any hope for those dealing with the software programs currently in place and not working with highly integrated enterprise applications but the more common mix between all kinds of small programs, Excel sheets, and homemade databases?

The good news is: yes, there is hope! A lot can be done to overcome the small and not-so-small inefficiencies by applying some simple and inexpensive rules and by asking the right questions.

UNDERSTAND YOUR REQUIREMENTS

What may appear as a platitude at the first glance is indeed one key element for the efficient and effective use of applications and data. What data needs to be actually stored, what is required to run the process (e.g. to order animals), what is required for legal or

compliance purposes, and what contributes to the research task?

A very simple example for illustration (certainly you can easily find a better one within your organisation): In your environment, is it necessary for any reason—progress, regulatory, or scientific—to document the actual room in which a particular experiment has taken place?

No? Fine, then you shouldn't bother your colleagues with recording that piece of information and your systems don't need to provide a functionality to capture it.

Yes? In this case it is clear that your system (again, these rules work for simple spreadsheets as well as for big enterprise-scale systems) should provide the capability to store that information.

UNDERSTAND YOUR SEARCH REQUIREMENTS

Again this looks like a very easy question and quite often the generic answer is that everything should be searchable. But actually the question isn't so easy and a wrong answer can have a lot of impact on the data search capabilities.

Coming back to our example where the room has to be recorded for any particular experiment: Is that done just as complementary information?

If yes, then it's probably acceptable that your system accepts any text value like "room 1", "R1", or "room/1" since they all fulfill that documentation requirement and are perfectly understandable for a human reader.

But what if one wants to list all experiments conducted in a particular room? In this case our pragmatic hands-on approach won't work. We simply wouldn't be able to run a search against all these free-text values reliably returning all the experiments we are looking for. Instead we would again have to screen all records manually.

And if "room 1" exists in multiple buildings, things get even worse.

If there is a need for structured search for a piece of information, you have identified a candidate for the next task.

UNDERSTAND THE ENTITIES

What are the information objects, also known as entities, to which data needs to be stored and processed?

How are they defined and, first and foremost, are they uniquely identified?

In our example, agreeing on a definition and common understanding is of course very easy. Everybody will understand and know what "room" means.

But we don't have to pick a complex research topic

to find examples where common sense alone won't do the job. What about terms like "project", "study", or "experiment"? Ask two members from different teams in any organisation about these quite common terms and you'll be surprised what level of confusion synonyms or (worse) homonyms can create.

It's absolutely key that all stakeholders using the systems and the data should have a clear understanding

of the definition of the entities and should be using the same name for the same things in order to allow structuring, searching, and sharing data.

...scientific results and measurements and the appropriate units need to be well defined within and across research groups as well as systems.

DEFINE RESULT TYPES AND UNITS

Actually there shouldn't be a need to mention it, but scientific results and measurements and the appropriate units need to be well defined within and across research groups as well as systems. Exactly the same name should be used if a result is measured or calculated the same way—and vice versa. And of course the unit should always be stored together with the value even in cases where a standardised unit has been agreed—too often terrible misinterpretations occur from comparing apples with oranges.

UNIQUELY IDENTIFY ENTITIES

This could be seen as part of understanding the entities, but since it's really crucial, it will get a separate paragraph. How are your entities uniquely identified?

Stressing again our little example: "Room 1" could work well as the identifier as long as there are no other buildings which also have a room with that name. In this case, only a combination of building and room would help to clearly link an experiment to its exact location.

And of course, exactly this unique identifier must be used in all systems and all situations where the room is referenced—any variation like "Room.1" or "Rom01" would break the link and make your efforts obsolete.

In database systems, often technical IDs are used for that purpose, but even in an Excel spreadsheet



SAFE

Scientific Animal Food & Engineering

www.safe-diets.com

there are mechanisms to make sure that only values from a defined list can be picked to indicate a reference.

DEFINE LEADING SYSTEMS

The concepts of entities and unique identifiers for them also can help to avoid another common issue: redundancy!

The attribute information related to an entity should appear only once in one system only and all other systems should only reference to this location.

Again our example: Of course the entity "room" can be referenced from diverse systems. The animal registry refers to it as the animal room, the list of conducted experiments links experiments to the procedure room and the homegrown equipment database shows the location of each piece of equipment. But the details to each room, security and sanitary status, size and capacity, maybe the historic room records should be stored in one system only.

If all of this homework is completed, the most important step needs to be done:

APPLY!

The described set of rules appears small and simple and can be applied to any kind and size of organisation and systems. Nevertheless, a lot of discipline and energy is required to implement it within a given data management environment and since these environments are evolving along with your business, this is not a one-time effort.

And of course applying these rules won't solve all issues at once and unfortunately it will not automatically reduce all of your reporting efforts to the click of a button.

But it will enable you to link bits and pieces together, to communicate and to share data and therefore it is a key to unlock the hidden value stored in all your spreadsheets, databases, and applications.

References

1. Wikipedia, http://en.wikipedia.org/wiki/Data_management.
2. Wikipedia, http://en.wikipedia.org/wiki/System_integration.

Andreas Staubi is CEO of a-tune software ag (www.a-tune.com), provider of data management consulting services for R&D organisations and of tick@lab, the management software for animals and animal facilities.

The diet and bedding solution for the security of your research



the best certified origin
 constancy
 dedicated skills and facilities
 diet and health safety
 total traceability
 control



Photo credits: Shutterstock.com, Fotolia.com

Route de Saint Bris - 89290 Augy - France
 Tel. :+33 (0)3 86 53 76 90- Fax : +33 (0)3 86 53 35 96
info@safe.evls.net