EQIPD – GA 777364
IMI2-2016-09-03

European Quality In Preclinical Data

WP5 – New quality management system in emerging and classical technologies to improve preclinical robustness

D5.3 Quality reporting module in research software (prototype)

Lead contributor	Lucas Noldus (18– Noldus)
	Lucas.Noldus@noldus.nl
Lead contributor	Björn Gerlach (2 – PAASP)
	bjoern.gerlach@paasp.net
Other contributors	Anton Bespalov (2 – PAASP)
	Arno van Gijssel (18 – Noldus)
	Jantine Duit (18– Noldus)
	Jeroen Kemerink (18 – Noldus)
	Tom Lavrijssen (19 – Janssen)
	Sami Virtanen (23 – ORION)
	Lee Monk (28 - UCB)
	Kenneth Drake (Stakeholder – Seralogix)
	Frank Thadewaldt (Stakeholder – enso GmbH)
	Nathalie Percie du Sert (Stakeholder – NC3Rs)
	Ben Fitzpatrick (Stakeholder – Tempest)
	William Buras (Stakeholder – Tietronix)
	John Kroehle (Stakeholder – DSI)

Gerry Herrera (Stakeholder – MED associates)

Document History

Version	Date	Description
V1.0	31 Mar 2020	First draft by BG
V1.1	22 Apr 2020	Updated draft by LN and BG Send out to all contributors
V1.1	28 Apr 2020	Report accepted by all contributors
V1.1	29 Apr 2020	Final Version

Abstract

Data transfer at the different stages along the research chain of study planning, data generation and the reporting of results is often prone to errors due to the different interfaces. This task aimed to focus on the transfer of information from the office to the lab (e.g. study design data) and experimental data from the lab to the office (e.g. acquired research data). For this purpose, a data transfer protocol was developed to facilitate the coding of study design and experimental data within the information-generating software and decoding at the side of data processing. Such intermediator “IT-protocols” are called CODECs (COding / DECoding) and was developed within this task. Several fixed parameters are described as a consensus from the working group and the format of the CODEC is described in a json-file format. A market analysis performed by Noldus investigated the need for the implementation of this CODEC. 230 responses were analysed showing that for 32% such an CODEC implementation is an unmet need for the software applications in use.

Methods

A group of different stakeholders was assembled to discuss the framework for the CODEC. This group consisted A) of software developers producing tools for the experimental design of studies and storage of data (Seralogix, Tempest, NC3R, enso) and B) of software developers producing software for data acquisition (Noldus, Med Associates, DSI). This group identified an essential set of parameters based on information obtained from WP2 and experience by members of the stakeholder group. The essential set of parameters was formatted according to ISO8601 and SEND (Standard for Exchange of non-clinical data) if there was an overlap between the standards and the identified parameters by the working group. A suitable file format for the CODEC was identified by the working group and agreed on. Noldus developed the first draft of the CODEC which was discussed with the other stakeholders and developed into the prototype. Finally, Noldus performed a market analysis to judge whether and how much resources shall be put into the implementation of this CODEC into the Noldus software.

Results

Development and description of the CODEC

The transfer of the research data from one software application to another requires a file format that can be used by each program. Hence, it was important to agree on a format that is versatile, interoperable and secure. The JSON file format seemed to be appropriate for the purpose of this CODEC. On that basis a prototype for the CODEC was prepared which can be found in Annex 1.

Beside the technical definitions, several experimental parameters had to be defined to provide universal reference points. In general, the values are separated into the experimental header and the trial data. For both mandatory and optional values exist, whereas the trial data is supplemented with a user specific category being optional as well. For each category, several specific data fields are defined together with the data type to be entered and an example (Table 1).

Table 1: Parameters for the EQIPD CODEC

	DATA FIELD	DATA TYPE	EXAMPLE
EXPERIMENT HEADER (mandatory)	ResearcherID	Text	"John Denver"
	Paradigm_name	Text	"Sociability Chamber"
	StudyID	Text	"K2TA78"
	ProtocolID	Text	"SC_12"
EXPERIMENT HEADER (optional)	Total_No_animals	Numerical	"48"
	Total_No_groups	Numerical	"4"
	Sample_Size	Numerical	"12"
	Power_cal	Numerical	".80"
	Effect_Size	Numerical	".33"
	Hypothesis	Text	"Cognitive dysfunction is connected the permanent hypofunction of the NMDAR"
	Comments	Text	"This is the last study in Q3 2018"

TRIAL (mandatory)	TestID	Text	"T001", "T002", …
	Randomisation_No	Numerical	"1", "19", "22", …
	Block_No	Text	"A", "B", "C", ...
TRIAL (optional)	TestgroupID	Numerical	"1", "2", "3", …
	Startdatetime_ timewindow	Date time (dd.mm.yy_hh:mm:ss)	"28.04.18_14:30:00"
	Enddatetime_ timewindow	Date time (dd.mm.yy_hh:mm:ss)	"28.04.18_14:30:00"
TRIAL (user-defined; *n)	IV_Label	Text	"Dosage", "Weight", "Genotype", etc
	IV_Type	Text	"Date", "time", "integer", "text", "boolean", "categorical", "numerical"
	IV_Format	Text	"dd.mm.yy", "hh:mm:ss", "null", "null", "null", "null", "x.xxxxx"
	IV_Physical dimension	Text	"nmol/l", "gram"

Market analysis for implementation of CODEC

In order to assess the need for implementation of the CODEC into a tool to automatically transfer data from an experiment design tool to a behavioural data acquisition system (such as EthoVision XT), Noldus Information Technology carried out market research in January-February 2020. Analysis and reporting were completed in March 2020.

A questionnaire was sent to 1562 users/prospects of EthoVision XT and 103 members of the EQIPD stakeholder group, resulting in 230 responses (13.8%), which is considered very high for this type of surveys. The results can be summarized as follows:

- Most respondents are active in behavioural neuroscience (60%) and based in North America (49%) or Europe (35%). Most (61%) use a generic tool to make experiment designs, with Microsoft Excel being by far most popular.
- For data acquisition (e.g. video tracking), 67% of the respondents use a Noldus tool. However, 45% of the respondents (also) use systems from other manufacturers, so the results can be generalized quite well.
- An Ulwick analysis showed that for 32% of the respondents, automatic data transfer from their experiment design tool to their measurement system is an unmet need. This is a very low percentage compared to other functionalities evaluated in user surveys in the past.

The report of the market analysis can be found in Annex 2.

Discussion

A functional CODEC was created and approved by stakeholders of this Task. This CODEC will be made available on the EQIPD web page to be available for further discussions. Additional publication channels are explored.

The market analysis showed a need for such a CODEC developed by EQIPD since many researchers use Excel as a tool to design studies. After designing the study, they most likely either copy and paste or even manually re-enter the information into the data acquisition software. However, despite this manual and time-consuming process, researchers do not see the need to facilitate this process with an automated approach. This is surprising because an automated approach could save time and reduce the errors while transferring the information.

Conclusion

Based on this result, and the fact that designing an automatic experiment design transfer tool from a free-form tool as Excel is a major undertaking, we recommend not to invest in this development at this stage. It is better to focus on getting the QMS and dedicated experiment design tools accepted in the community, and once that has been achieved, develop automatic interfaces from a specific experiment design tool to the most common behavioural and physiological testing systems.

ANNEX 1

The EQIPD CODEC with mock values:


```
{
   "Experiment header / mandatory" : {
      "ResearcherID" : "John Denver",
      "Paradigm_name" : "Sociability Chamber",
      "StudyID" : "K2TA78",
      "ProtocolID" : "SC_12"
   },
   "Experiment header / optional" : {
      "Total_No_animals" : "48",
      "Total_No_groups" : "4",
      "Sample_Size" : "12",
      "Power_cal" : ".80",
      "Effect_Size" : ".33",
      "Hypothesis" : "Cognitive dysfunction is connected the permanent hypofunction of the NMDAR",
      "Comments" : "This is the last study in Q3 2018"
   },
   "Trial / mandatory" : [
      {
         "Label" : "TestID",
         "Type" : "integer",
         "Unit" : null
      },
      {
         "Label" : "Randomisation_No",
         "Type" : "integer",
         "Unit" : null
      },
      {
         "Label" : "Block_No",
         "Type" : "integer",
         "Unit" : null
      }
   ],
   "Trial / optional" : [
      {
         "Label" : "TestgroupID",
         "Type" : "integer",
         "Unit" : null
      },
      {
         "Label" : "Startdatetime_timewindow",
         "Type" : "date time",
         "Unit" : null
      },
      {
         "Label" : "Enddatetime_timewindow",
         "Type" : "date time",
         "Unit" : null
      }
   ],
   "Trial / user-defined" : [
      {
         "Label" : "Dosage",
         "Type" : "numerical",
         "Unit" : "nmol/l"
```

```
      },
      {
         "Label" : "Weight",
         "Type" : "numerical",
         "Unit" : "gram"
      },
      {
         "Label" : "Genotype",
         "Type" : "numerical",
         "Unit" : null
      }
   ],
   "Data" : [
      {
         "TestID" : "T001",
         "Randomisation_No" : "1",
         "Block_No" : "A",
         "TestgroupID" : "1",
         "Startdatetime_timewindow" : "2018-10-15T11:44:01.000000+02:00",
         "Enddatetime_timewindow" : "2018-10-15T11:45:35.000000+02:00",
         "Dosage" : "3.1",
         "Weight" : "25",
         "Genotype" : "12"
      },
      {
         "TestID" : "T002",
         "Randomisation_No" : "19",
         "Block_No" : "B",
         "TestgroupID" : "2",
         "Startdatetime_timewindow" : "2018-10-15T11:51:33.000000+02:00",
         "Enddatetime_timewindow" : "2018-10-15T11:53:21.000000+02:00",
         "Dosage" : "1.5",
         "Weight" : "24",
         "Genotype" : "12"
      }
   ]
}
```