Paper PP14

A Two-Step Procedure – Select and Act – How Does It Work?

Catharina Dahlbo, Capish, Malmö, Sweden
Eva Kelty, Capish, Malmö, Sweden

ABSTRACT

While an ontology-based information model and a property graph database establish the backbone of a data platform for the future, also a well-designed user interface that facilitates the "communication" between end-user and machine is needed. A powerful point-and-click user interface has been developed where the platform takes advantage of advanced data storage and retrieval. Briefly, data from an information object, a Holon, can be displayed in e.g. listboxes, graphs, tables, and searches where the end user can see and work with all the unique values. By making selections in such a component, the next click opens a menu displaying the possible actions that can be taken in the context of the current selection. This special feature lays the foundation for the user of being able to ask more complicated questions.

This poster will show how this two-step procedure, select-and-act, allows the user to perform different complicated actions very quickly.

INTRODUCTION

Often, data marts are built every time when data needs to be evaluated in a new context. In most cases this means that information is lost as not all data is integrated in the individual data marts due to the complexity of creating appropriate information models. At the same time, querying and analyzing data sets in existing solutions for data exploration and visualization are often limited to either flexibility or user-friendliness. They are also commonly sensitive to structural changes of the data sets.

An alternative to these data handling processes could be a data repository that enables instant access and reuse of all available data. With these features new hypotheses can be checked without time consuming data extraction, transformation, and load procedures. But the utility is not perfect until the repository is combined with an interface that connects the end users to the data to save even more time.

Capish Reflect leverages modern web technologies to provide an intuitive user interface connecting the end user to the data store using standard web technologies. It is built on an ontology-based information model using graph database technology for data storage.

In order to provide a powerful and flexible solution with new opportunities it is based on a combination of different useful components:

- Information Model - How the data is modelled [1]
- Ontology - How the data in the model is described [2]
- Graph Database - How the data is stored [3]
- Reflective Logic - How the graph database is queried and searched [4, 5]
- Interface – How the data is explored and visualized to the end user

Capish Reflect is built on a new paradigm on how to effectively query and retrieve information from a data set. It exploits the existence of direct and indirect relationships in the information model. As an example, a search for a value could potentially not only retrieve all the information objects, Capish Holons [6], containing that value but could also be configured to retrieve all directly or indirectly related information.

Utilizing a web-based interface based on a graph database containing curated and integrated data gives great flexibility in data exploration and visualization in many ways. Here we will focus on one aspect, i.e. various ways of selecting data and taking action.

SELECT & ACT

The special feature, select & act, is a two-step procedure that lays the foundation of what could be possible to do in the next step, such as ask more complicated questions, work with complex cohorts or find previously unknown relationships. Briefly, data from a Capish Holon can be displayed in e.g. listboxes, graphs, charts, tables, and searches where the end-user can see and work with all the unique values. By making selections in such a component, the next click opens a menu displaying the possible actions that can be taken in the context of the current selection such as filtering, highlighting, adding to list of members, and showing selected Capish Holons.

SELECTION

The purpose of a selection is to identify and define what you later want taking action on. A selection criterion is defined based on the selected values. This criterion is then used as a query when taking action, e.g. creating a filter or highlight.

A selection identifies values of interest and does not in itself make anything happen.

There are different ways of marking values of interest when doing a selection, mouse clicking and brushing.

BY MOUSE CLICKING

1. Select value from any field

The interface allows to do a free text search for a specified text in the whole database. All individual hits, regardless of which Holon it belongs to, are listed after a search. The values are distinctly displayed, and one value can be selected by a mouse click and more values can be added to the selection by using Ctrl click. In the example below the search for the value abnormal gives hits in different areas such as Blood Analysis and ECG and there within different fields. One or more values can be selected for further actions depending on the question.

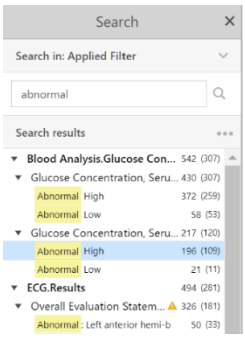
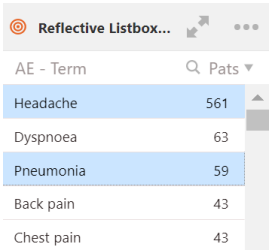


Figure 1. An example of a free text search in the whole database. The search is for abnormal and the hit Abnormal High in Glucose Concentration can then be selected (marked) by a mouse click.

2. Select or deselect values from the same field

The same way of selection, by mouse clicking, can also be done in e.g. a component called listbox. A listbox differ from the free text search that it is only values within the same field that is shown. By using Ctrl/ALT click multiple values of a field can be combined with OR, AND or AND NOT.

OR is used as default when selecting multiple values in a field. In the example in the figure we want to look for *Patients that have had Headache OR Pneumonia or both Headache AND Pneumonia*.



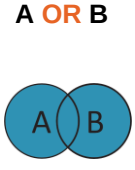
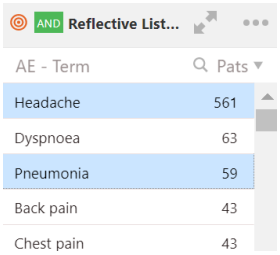


Figure 2. Both Headache and Pneumonia are selected (marked) by mouse clicks and the operator OR is used as default between the values within the field AE - Term.

In the settings for operators, OR can be changed to AND (green marked area) in a listbox like the one in figure 2 and 3. The operator AND is used for selection when multiple values of a field should be fulfilled. Here we want a more specific search for *Patients that have had both Headache AND Pneumonia*.



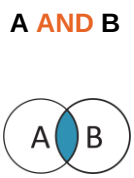
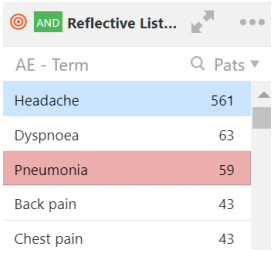


Figure 3. Changing the operator from OR to AND allows to combine values within a field with AND to get a more specific search.

By combining Ctrl click + Alt click, NOT can be added to the selection with AND which help us to exclude values. Here we want to search for *Patients that have had Headache AND NOT Pneumonia***.** The selection can contain any number of inclusions and exclusions.



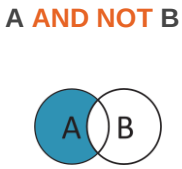


Figure 4. With the operator set to AND the possibility to exclude values increases the number of combinations further.

Now we have looked at values from the same field but values from different fields can also be selected and combined with OR, AND and AND NOT, if the fields are collected in a multibox, a box containing different listboxes.

BY BRUSHING

In a chart, values are selected by brushing the area of interest, it could be one or more values in the selection. In the chart in figure 5, the area to the right is brushed and includes values of interest.

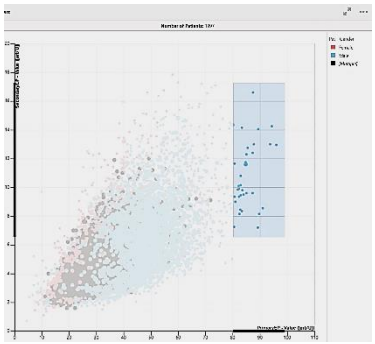


Figure 5. In the scatter plot the blue area is selected (marked) by brushing and contains values of interest.

THE ACTION MENU

Once the data of interest has been selected, several actions are available in an action menu belonging to the selected component (listbox, scatter plot, table etc), see figure 6. The actions menus may differ slightly, depending on component.

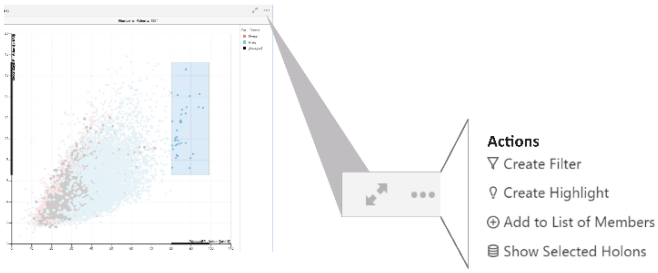


Figure 6. The action menu is found in the upper right corner in all components.

The actions in the action menu are independent of each other and multiple actions can be used for the same selection. The outcome of an action varies depending on in which order the actions are taken, and that depends on the question. A valuable feature in the application is that an action can be done both on values and on patients which also results in different outcomes.

Any action we choose to take, impacts overall in the application. This makes it easy to see values in the context of the whole data set.

The most common actions are shown in figure 6 above and will be described in the next section.

ACT ON SELECTION

ADD TO LIST OF MEMBERS

When values have been selected, cohorts can be created with e.g. patients having these values. By adding the selection to List of Members, all patients in the database, meeting the criteria and having the selected values, are collected in a cohort. More patients can be added to the cohort by making further selections or handpicking members at any time. Patients can be removed from the list as well.

Once the list exists you can choose to look further into each patient's data, or you can use the list as a selection and take more actions. A cohort can later be further compared to other cohorts since it is possible to save multiple Lists of Members.

Since end-users might be interested to evaluate the data from different perspectives, it is also possible to define lists with different vantage points. By adding studies to a list, it is possible to see from a clinical study perspective. Finding patients that seem to be similar to each other is another feature available in the interface. These similar patients could also be collected in a cohort.

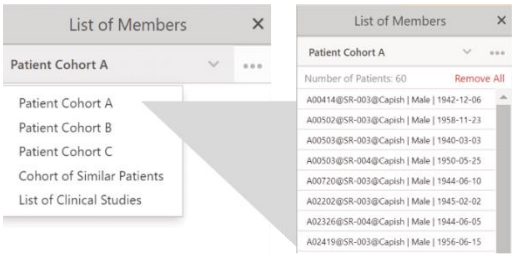


Figure 7. The List of Members view to the left contains 5 different cohorts where patients/studies can be collected. The list to the right shows some of the patients collected in Cohort A.

HIGHLIGHT

When doing a highlight, a selection can be shown in context of all data. Everywhere in the application the values are marked as yellow and so are other also values related to the patients having the selected values.

Creating a highlight directly on selected values will also highlight other related values occurred for the patients having the selected ones. For example, all values occurred at the same time as the selected values are then highlighted. By adding a selection to a 'List of Members' first and then highlighting the patients in the list will highlight all values ever occurred for the listed members throughout the application. This means that the difference between doing a highlight on values or members first is that a broader result set is obtained with the member way. This approach is beneficial when searching for all related data inclusive data of the same type such as all adverse events ever happened for the patient.

Figure 8 shows a highlight of selected values and figure 9 shows a highlight on values related to the patients having the selected values in figure 8. Here we can see in figure 9 that not only the selected values are highlighted but also all other values for the selected patients for the same type of measurements.

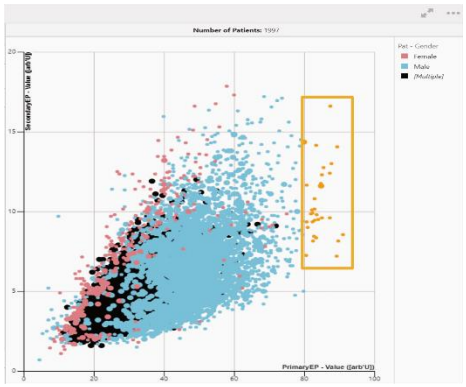


Figure 8. The selected values are highlighted when only doing a highlight on these values.

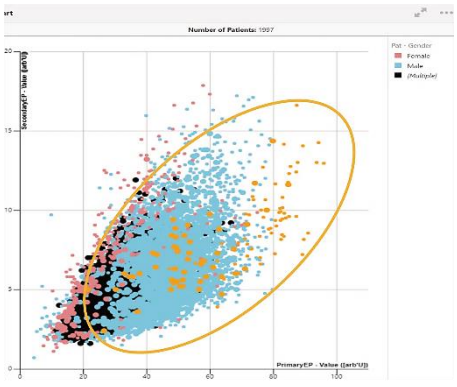


Figure 9. By doing a highlight on patients having the selected values in figure 8, all values ever happened for the patients are highlighted.

Figure 10 shows how it could look in another part of the application when a highlight is activated. In this part we get a very quick overview in which treatment arms, studies, age span etcetera the patients belong to and which patients that are included in the action. Changing to another tab, gives information of other parts of the data e.g. medications, adverse events, lab values etcetera for these patients.

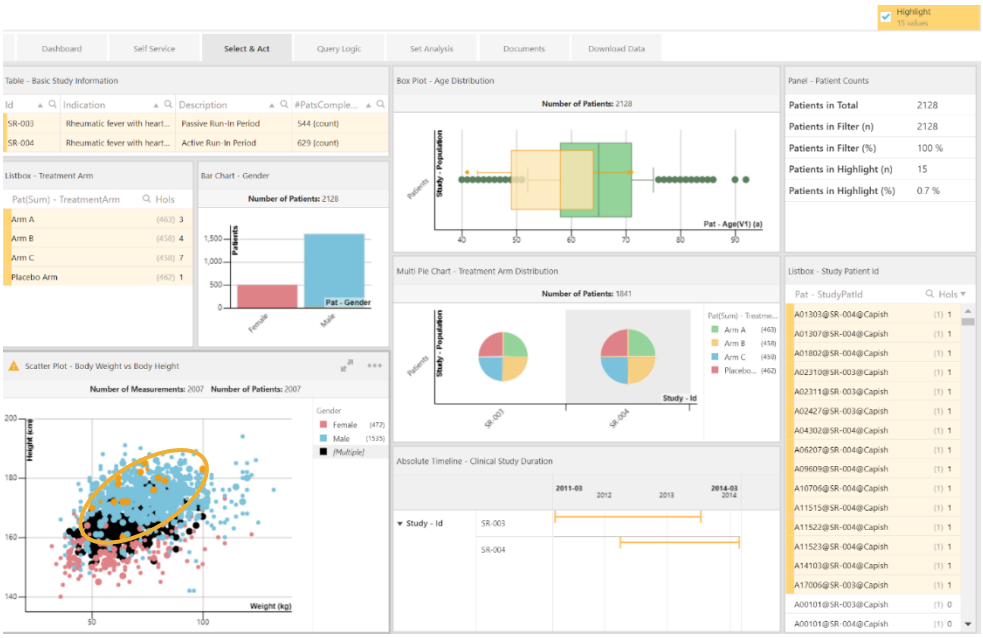


Figure 10. In another part of the application, data related to the patients having the selected values are highlighted.

FILTER

By using filter on the selected values, the data set is reduced to only contain the selected values and values related to them in the whole application, independent of component. Only patients having the selected values are left in the reduced data set.

Adding a selection to a List of Members first and then filtering the data set is reduced to contain all values related to the members, e.g. patients. Unlike highlighting, the data set is reduced and is so in the whole application. A filter can be created from a highlighted selection and thereby create a reduced data set.

The figure 11 below shows a reduced data set that contains all the values related to patients having the selected values in figure 5, thus the values highlighted in the figure 9.

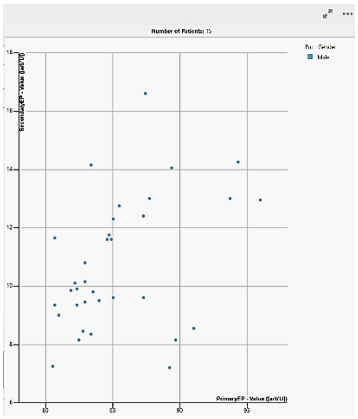


Figure 11. Here is a reduced data set containing all related values to patients having the selected values in figure 5. The values are the same as were highlighted in figure 9.

Figure 12 shows an overview from another part of the application and where only the filtered data is included in the visualizations. For example, at a quick glance in the pie chart we can see that there are only males and 6 of them belongs to Arm C in study SR-004.

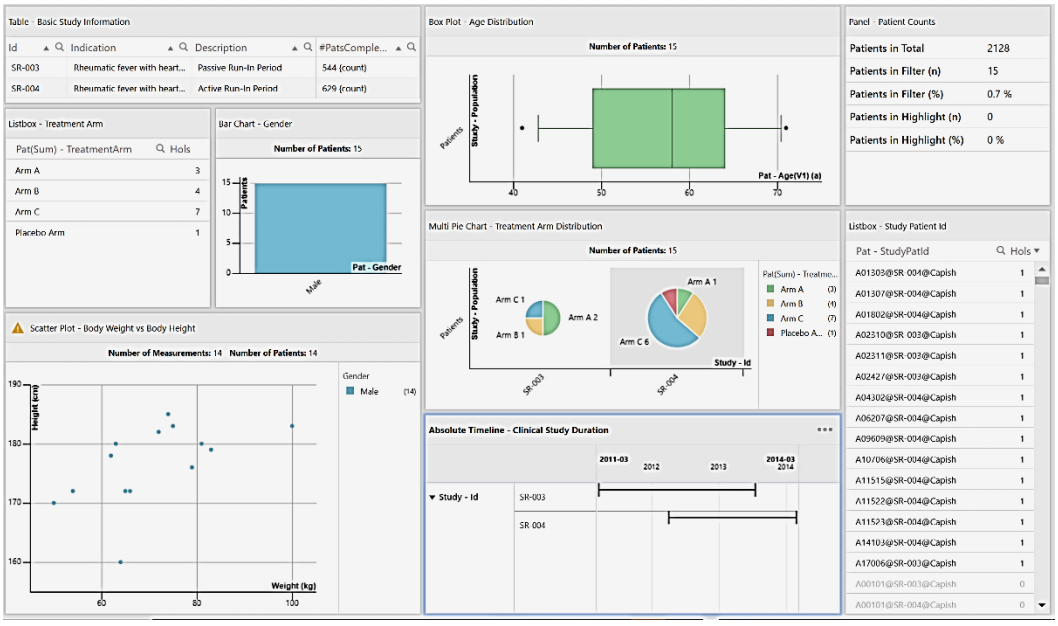


Figure 12. The difference between this figure and figure 10 is that the data set here is reduced to only contain the reduced data.

SHOW SELECTED HOLONS

Looking deeper at selected values and other related information is possible by doing the action ‘Show Selected Holons’.

The database view of the selected data is shown, and it is also possible to navigate further into the database to look at both data and metadata.

In the picture below, the Capish Holons including the selected values in a component are listed to the upper left. By selecting a specific Capish Holon type to the left, a list of Capish Holons that are of that Holon type are visible to the right. Clicking on a Capish Holon will make details visible for the selected Capish Holon in the lower right. The graph to the lower left shows data related to the current Capish Holon. Navigating in the graph moves the user to related values.

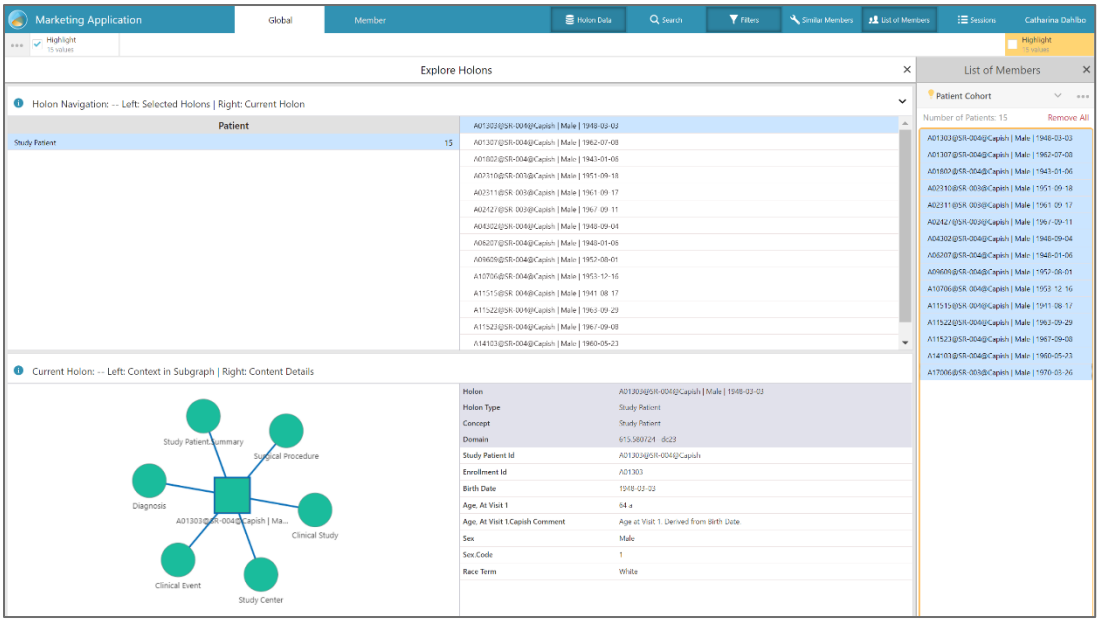


Figure 13. More details about the data are available when exploring data and its metadata in the database via Show Selected Holon.

CONCLUSION

By separating the select and act functions the user’s flexibility vastly increases. The user can for example choose between reducing the data set based on the selection or explore the selected data in context through a highlight. Also, it lets the user easily build cohorts for further exploration as well as study each selected patient in more detail. The benefit of this approach does not lie in the individual steps alone but rather in the total concept. The strength is that the search engine only needs to keep track of a finite amount of Capish Holons and their relations which enables the high search speeds necessary for presenting an interactive user-friendly interface, while still leaving lots of information at the fingertips of the end-users.

REFERENCES

1. E.Kelty, “DV07 - Effective Data Modelling for Effective Data Visualization”, presented at PhUSE EU Connect 2018
2. P.Tormay and A. Berg, “DH07 - Harness the power of ontologies to build better information platforms”, presented at PhUSE US Connect 2019
3. A. Berg, H. Drews and C.Dahlbo, “*PP05 – New Approach to Graph Databases*”, presented at PhUSE EU Connect 2018
4. C. Dahlbo, A. Workneh and H. Drews, *“PP04 – Unique Technology for the Future”*, presented at the PhUSE EU Connect, 2018.
5. S. Gestrelius and H. Drews, *“Patent US20130246438 –Reflective logic unlocks knowledge in datasets”*, 2013
6. C.Dahlbo and E.Kelty, *“PP20 - Holons – What, Why and How”*, ”, presented at PhUSE EU Connect 2019

ACKNOWLEDGMENTS

We want to thank Patric Moreau for his contribution to the virtual poster session.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Catharina Dahlbo
Capish Nordic AB
Lokgatan 8
211 20 MALMÖ, Sweden

+46 (0)40 10 88 80
catharina.dahlbo@capish.com

Eva Kelty
Capish Nordic AB
Lokgatan 8
211 20 MALMÖ, Sweden

+46 (0)40 10 88 80
eva.kelty@capish.com

www.capish.com