Data-driven indicator classification and selection for dynamic dashboards: The case of Spanish universities

Rafael Molina-Carmona\textsuperscript{1}, Faraón Llorens-Largo\textsuperscript{2}, Antonio Fernández-Martínez\textsuperscript{3}

\textsuperscript{1} Cátedra Santander-UA de Transformación Digital, University of Alicante, Ctra. San Vicente s/n 03690, Alicante, Spain, rmolina@ua.es
\textsuperscript{2} Cátedra Santander-UA de Transformación Digital, University of Alicante, Ctra. San Vicente s/n 03690, Alicante, Spain, faraon.llorens@ua.es
\textsuperscript{3} Department of Computer Science, University of Almería, Crta. Sacramento s/n, La Cañada de San Urbano, 04120, Almería, Spain, afm@ual.es

Keywords
KPI classification, KPI selection, Dashboard design, Data-driven design, Spanish universities KPI.

1. ABSTRACT

In the context of business, dashboards are visual tools that display the most important information about the organization needed to help the top management to make decisions. Since it is important to just provide the relevant and objective-oriented information, the number of indicators included in the dashboard must be kept at minimum. Therefore, the crucial aspect when designing dashboards is the selection of the suitable Key Performance Indicators. To help to carry out this task, we propose a classification and selection methodology, based on the values of the own indicators. This methodology is performed in two main steps: the classification of the indicators in three categories (emergent, hot and consolidated indicators) and the selection of the suitable set of KPIs based on the organization strategies. To illustrate the application of this methodology, we also present a practical case of indicator classification and selection for Spanish universities based on the extensive UNIVERSITIC report.

2. INTRODUCTION

A dashboard is a business tool that displays a set of indicators and other relevant information to a business user. The information is usually represented graphically and must include the Key Performance Indicators (KPIs) involved in achieving the business objectives. This way, the dashboard is oriented towards decision making to optimize the organization's strategy.

A dashboard should be displayed and understood at a glance, so it should include only those KPIs needed to represent the main objectives and strategies that top managers need to make decisions. This means that the number of KPIs in the dashboards is usually very small and represents abbreviated information in the form of summaries or exceptions. In the case of Information Technologies (IT) and universities, the dashboards end up containing indicators related to very specific aspects such as personnel or budget.

Dashboards are sometimes conceived just as fixed pictures of the KPIs that most influence the achievement of business objectives. In some other cases, dashboard data are displayed in real-time after retrieval from one or more data sources or even are interactive, allowing the navigation into particular aspects of the display or switch between facets or views of the data. However, they seldom allow the selection of different KPIs. In summary, dashboards are usually dynamic in the way information is displayed, but not in the selection of KPIs.

Furthermore, many data and indicators are usually collected and developed in organisations, although only a few end up being part of the KPIs of the dashboard. The selection of these KPIs has been the subject of numerous studies that highlight the complexity of this selection and the need for it to be ad-hoc designed for each organization according to its characteristics.
Our proposal is to use the values of the indicators to classify them and determine which of them are most suitable to be part of the dashboard, in what we have called data-driven indicator classification and selection. This way, decision making takes place at two levels: on the one hand the values of indicators and their evolution help the dashboard designer to classify and select the KPIs that will be part of this dashboard, and on the other hand the KPIs themselves will help the top management to make their business decisions. Moreover, there is a second derivative from this proposal: the data themselves will tell us how the different indicators evolve and, therefore, when the KPIs of the dashboard are likely to be replaced by more significant ones. In short, more dynamic dashboards will be obtained, to be adapted to the changing environment conditions.

The document is organized as follows: in section 3 we review some previous concepts and work related to dashboard design and indicator selection. In section 4 we present our data-driven indicator classification and selection proposal, explaining the methodological aspects of the model. Section 5 is devoted to a practical example of the selection of indicators based on UNIVERSITIC report, which gathers the IT indicators of Spanish universities. Finally, the conclusions are set out in section 6.

3. BACKGROUND

3.1. Dashboard design

All organizations need an information system that allows the communication of key strategies and goals and the decision-making. It is what Eckerson (2010) calls the “organizational magnifying glass”. This author considers that the dashboard is the organizational magnifying glass that translates the organization’s strategy into objectives, metrics, initiatives and tasks.

More concretely, Few (2006) considers that “a dashboard is a visual display of the most important information needed to achieve one or more objectives; consolidated and arranged on a single screen so the information can be monitored at a glance.” In few words, a dashboard must contain:

- Visual information
- Important information
- Objective-oriented information
- Little information
- Understandable information

The main objective of a dashboard is transforming data into information and turn this into knowledge for the business. More precisely, for Eckerson (2010), the goals of a dashboard are:

- Monitor critical business processes and activities using metrics of business performance that trigger alerts when potential problems arise.
- Analyse the root cause of problems by exploring relevant and timely information from multiple perspectives and at various levels of detail.
- Manage people and processes to improve decisions, optimize performance, and steer the organization in the right direction.

It is important to note that the dashboard is useful to make decisions that optimize the organization’s strategy, but not to study in detail certain issues, cross variables, or develop concrete actions.

Few (2006) describes an interesting set of features for the dashboards:

- A dashboard is a visual display, so the information is presented as a combination of text and graphics (charts, grids, gauges, maps...), but with an emphasis on graphics. An efficient and attractive graphical presentation can communicate with greater efficiency and richer meaning than text alone.
- A dashboard displays the information needed to achieve specific objectives, so its design requires access to complex, unstructured and tacit information from diverse sources. The required information is often a set of KPIs, but other types of information might also be needed.
• A dashboard fits on a single computer screen, so that it can all be seen at a glance. No scroll or multiple screens should be allowed to allow a quick comprehension of the information.

• A dashboard presents updated information, so some indicators may require a refresh in real time but others may need to be updated with other frequencies.

• A dashboard is used to monitor information at a glance, so the information it presents is abbreviated in the form of summaries or exceptions. No details can be monitored at a glance. Its primary job is detecting that an action is required, but the action cannot be determined solely with the information contained in the dashboard.

• A dashboard has small, concise, clear, and intuitive display mechanisms. It has minimal unnecessary distractions that could cause confusion and the aspects of usability and meaning should centre its design.

• A dashboard must be customized, so that its information must be tailored specifically to the requirements of a given group or function.

3.2. Indicator selection

The information in a dashboard is a set of Key Performance Indicators. These indicators are usually high-level measurements of how well an organization is doing in achieving critical success factors. To determine what KPIs take part in the dashboard, the designer must consider the audience level in the management structure. Therefore the details should be removed as the target audience moves up in the management structure to avoid information overload. Moreover, a typical dashboard is usually made of a few KPIs, including only those that are strictly necessary (typically between 7 and 10).

Selecting the KPIs is not an easy task. In most cases this task is manual and specific to a particular case. However, there are some research projects that seek to formalize and automate all or part of the dashboard design process, from selecting KPIs to defining visualization tools. In this section we highlight some of the main contributions that have been made on this subject.

Chowdhary, Palpanas, Pinel, Chen and Wu (2006) propose an efficient and effective model-driven dashboard design technique. A model is a formal specification of the function, structure and behaviour of a system from a specific point of view, represented by a combination of drawings and text. The models are used as the primary source for selecting the KPIs and designing, constructing and deploying the dashboards.

Kuntz (2012) presented a semantic dashboard description language used in a process-oriented dashboard design methodology to help overcome known challenges of business process monitoring, such as the difficulty to build appropriate dashboards from complex data sources to best monitor given goals.

As regards methodologies for building dashboards, Brath and Peters (2004) advise following an iterative model of creating dashboards for better designs. The design iteration and the use of sketches and prototypes help identify the needs and requirements and refine vague design ideas into the best possible solution.

The selection of the KPIs must meet a number of constraints that we have already discussed: they must be directly related to the organization’s goals, they must focus on few key metrics, they must consider the state of the organization and be adapted to the business model and features. An interesting work is that of Keck and Ross (2014), that have investigated solutions to the selection of KPIs through the use of machine learning techniques in the particular case of a call centre. In this context of dynamism they have consider the problem as one of multi-label classification where the most relevant KPIs are labelled and selected later.

Our proposal is framed in this area, in the methodologies for classifying and selecting KPIs to design and build dashboards. Our final aim is to formally propose a data-driven KPI selection model that will produce more dynamic dashboards adapted to the needs of top management.
4. DATA-DRIVEN INDICATOR SELECTION

The concept of Data-Driven Design comes from the field of digital design (Harris, 2013), although it has also been applied in videogames, apps and industrial product design. This paradigm advocates for tailoring products to user’s preferences, goals, and behaviours. The collection of data from the users is, therefore, the starting point for every design. The concept of Simulation-Driven Design, although it is closely related of that of Data-Driven Design, comes from the field of industry. Simulation-Driven Design is a development process wherein numerical simulation is used to capture all aspects of a system, and modifications to that system are based upon the results derived from a numerical model (Karlberg, Lofstrand, Sandberg, & Lundin, 2013; Miller, 2004). Both paradigms, Data and Simulation Driven Design, imply the shift from top-down design models to data-based methods, where the data come from what users want and need or from numerical simulations of the industrial pieces.

In the field of business, Data-Driven Decision Support systems make up a category of decision-making systems that emphasizes access to and manipulation of a time-series of internal organizational and external data. These systems range from simple file systems accessed by query and retrieval tools, to data warehouse systems that allow the manipulation of data by computerized tools and On-line Analytical Processing (OLAP) that provide the highest level of functionality and decision support (Power, 2008). From this point of view, using data-fed dashboards in the organization obviously implies that the decision-making system is data driven. In our case, however, the nuance is different. We are not just talking about data-driven decision-making but about the use of data to determine which KPIs are part of the dashboard.

Before describing our proposal on how to classify and select KPIs, it is worth highlighting the main steps to design and deploy a dashboard in an organization. Schematically, these stages are presented in Figure 1.

![Figure 1. Stages for the design and deployment of a dashboard](https://via.placeholder.com/150)

Although in this paper we focus on the indicator selection stage, we briefly review what each step entails and then focus on the indicator selection stage:

- **Indicator generation.** In an organization, the amount of information handled is massive. Organizations are increasingly collecting and storing a vast amount of data from all their constituent units, which are often difficult to manage. These data are usually numerous, poorly organized and poorly formalized, and have a diverse nature. The first step in creating
a really useful dashboard is to be able to extract information from that data. To this end, the fundamental step is to convert these data into indicators. An indicator is a measurable variable used as a representation of an associated factor that can be non-measurable. Therefore, the key characteristics of the indicator are to be measurable and to have a meaning for the organization. The output of this stage is a list of indicators, that is, a list of variables that can be measured and have a meaning.

- **Indicator selection.** The indicators list from the previous stage is just a list of variables that contain information but that are difficult to transform into true knowledge for the organization due to their large amount and their heterogeneity. This stage is devoted to the classification and selection of indicators that make up the most suitable KPI set and it is the aim of this research. The details about this issue are provided in the next section.

- **Dashboard design.** Once the KPIs are selected, an effective way of presenting them must be designed. As stated in the previous sections, for the dashboard to be effective, a suitable combination of text and graphical representation is crucial. There is a wide research about this issue than can be of great help for designing the most attractive and effective dashboards.

- **Dashboard deployment.** The last stage is deploying the dashboard and monitoring its functioning. The right technologies should be selected and an iterative cycle of revisions is crucial to maintain the dashboard in the best conditions.

In the following section the indicator selection stage is described in detail.

4.1. **Indicator selection stage**

The information from the indicator generation stage is a list of indicators whose characteristics are:

- It is a very broad list, with a large number of indicators.
- Indicators come from different areas and units of the organisation.
- Indicators belong to very diverse levels of abstraction, and can represent very specific factors or be aggregated indicators that characterize a very general factor.
- Indicators have very different typologies and units of measurement.

The objective of this stage is to select the most suitable indicators to be part of the dashboard. For this purpose, we propose two steps: classification and selection.

The first step is to characterize the indicators and classify them into different categories. The criteria for classification will be based on two variables:

- The age of the indicator, i.e. the time that the indicator has been in the organization's catalogue.
- The evolution of the indicator over time, based on a defined date range.

In particular, we propose a classification into three categories, which will be the candidate indicators to be part of the dashboard. These categories are:

- **Emerging indicators:** These are usually recently incorporated indicators that still have a low value, but with considerable growth. It should be borne in mind that, when starting from such low values, any improvement, however small, produces considerable proportional growth, so growth must be viewed with caution.

- **Hot Indicators:** These are indicators that have an intermediate value, but follow a growing evolution. They are therefore indicators of issues that are still under development, which have intermediate values but are constantly growing.

- **Consolidated indicators:** These are very stable indicators, which already have a high value, so there is almost no space for growth.

In addition to these three categories, some others could be proposed, especially for those indicators with negative trends, which will be of great interest to the organisations.

The indicators can be represented graphically as in figure 2. The horizontal axis represents the moment when the measurement is made from the first time the indicator appears in the catalogue.
This time is measured in units corresponding to the update interval of the indicator. For example, if the indicator is updated and measured annually, the time unit will be the year. As for the vertical axis, it indicates the value of the indicator with respect to the maximum possible, that is, in percentage terms. In the case of absolute (non-percentage) value indicators, we must find a way to set a maximum and a minimum for their value (for example, by analysing the values of the indicator and taking the upper and lower extremes of the range of values as the maximum and the minimum achievable).

![Figure 2. Proposed classification for indicators](image)

Figure 2 shows three indicators each for one of the categories. Assuming that the time unit chosen for the update is the year, the evolution of these indicators over the last three years is represented. Indicator $I_1$ started to be collected one year before the first measurement and its evolution in these three years is increasing, although starting from very small values. This is an emerging indicator. Indicator $I_2$ had already been collected in the first measurement for two years. Its values are higher and, although it is increasing, it does not evolve so drastically. This is a hot indicator. Finally, the $I_3$ indicator had already been in the catalogue for 5 years in the first measurement considered. It has very high values and its evolution, although positive, is very slow. This is a consolidated indicator.

In figure 2 it is also possible to define roughly the bands in which the indicators of the different categories will appear. In addition, the slope of the curves of each indicator and its position in the band will give us additional information. Thus, on the one hand, the greater the (positive) slope of the curve, the greater its growth and the faster its evolution. On the other hand, the higher the curve is in its band, the greater its value but the less likely it is to continue growing.

Once the indicators have been classified, we propose that the selection be made according to the strategies adopted by the organization. For example, in the case of the Universities and the catalogue of UNIVERSITIC indicators maintained by Spanish universities (Gómez Ortega, 2016, 2017; Piriz Duran, 2015), their final objective is to serve as a guide for the improvement of our universities, which is why we place emphasis on novel aspects and implicitly assume the generalized ones. All in all, the most prominent hot and emerging indicators will be candidates to constitute the set of dashboard KPIs, while the consolidated indicators will be outside the critical set of KPIs.

5. THE CASE OF SPANISH UNIVERSITIES

UNIVERSITIC (Fernández Martínez, Llorens Largo, & Hontoria Hernández, 2015) is a catalogue with more than two hundred indicators of the Spanish university system, gathered by a specifically developed web application, called kTI (Fernández Martínez, Llorens Largo, & De Andrés, 2011). The indicators are structured in two layers (IT Description and IT Management), in 11 axes (5 in the IT
In this section we explain how we have made the selection of indicators based on the model presented in the previous section. To analyse its behaviour, we will look at the evolution in 2015, 2016 and 2017 (Gómez Ortega, 2016, 2017; Piriz Duran, 2015). The first aspect is the classification of indicators into the three chosen categories (emergent, hot and consolidated), presented in the next sections.

5.1. Emergent indicators

In this category, the most recent indicators with a low value but high growth in the three years considered have been classified. The emerging indicators, according to this criterion, are:

- **Virtual desktops**: Number of software configurations (different programs) offered in the virtual desktop/applications catalogue for teaching practices. It is an indicator that was incorporated for the first time in the 2015 report, practically doubling in three years, from just over thirty to almost sixty.

- **Massive Open Online Courses (MOOC)**: Number of MOOC courses in which the university actively participates (in exclusive or shared) and number of MOOC courses in which the university actively collaborates. These indicators have been incorporated for the first time in the 2015 report and have almost doubled in these three years. Currently, universities participate on average in 10.86 MOOC courses and collaborate in 7.39.

- **Interoperability**: Indicators related to facilitating interoperability and the use of shared resources seem to be starting to move. Although it is necessary to be cautious, since they start from very low values, therefore any small increase in absolute value represents a high increase when converting it to relative values. Taking into account the time perspective of these last three editions, the number of interoperability services used by the university has increased from five to seven, the number of interoperability services offered by the university has doubled (from one to two) and the number of institutions with which the university is linked through interoperability services has increased from ten to almost fifteen.

- **Open Government**: The indicators of the objective of exchanging information with other institutions in a rapid and efficient manner, incorporated for the first time in UNIVERSITIC 2015, have experienced spectacular growth. The number of universities that have a transparency portal has doubled in three years, being currently 90%, so we can consider it not only as emerging but also consolidated. The percentage of universities where there is an open data initiative or RISP plan has also doubled, although there is room for improvement, as it is around 20%. The number of published datasets has quadrupled, currently being 13.

5.2. Hot indicators

These indicators have an intermediate value but are evolving steadily. These issues are still under development. About half of the universities consider them and now they are being served by the rest.

- **Wireless network**: As expected, university Wi-Fi connections and the number of different users continue to grow. The indicators related to Wi-Fi (number of Wi-Fi connections per year, the average number of Wi-Fi connections per day and the number of different university students who connect to the Wi-Fi of the university annually) do not stop increasing report after report, placing us at present in almost twenty million annual connections, which represents more than fifty thousand daily connections and about forty thousand different users.

- **IT project portfolio**: The indicator "Do you annually design a well-defined IT project portfolio that is prioritized and approved by the university's governance team?" has a value close to 50%, but with a growth in universities that have participated in the last two editions
of 11%. In the previous edition (2016) it also had a growth. It has gone from 30% in 2015 to almost 50% today. It is therefore a hot topic, evolving favourably, but very slowly.

- **Information security:** Indicators aimed at measuring the goal of providing services with adequate safety conditions were introduced in the 2015 edition and have not stopped growing since then. The overall rate of maturity has gone from 24 to 39, the percentage of universities assigned the role of Information Officer has gone from 58% to 74%, those assigned the role of Service Officer has gone from 61% to 78% and those assigned the role of Security Officer has gone from 71% to 80%, hoping that in the near future they will be extended to practically all universities. It is therefore a matter of the utmost importance and concern.

- **Compliance with IT-related regulations:** The indicator "Internal control to ensure compliance with IT-related regulations" has experienced a 13% growth, reaching almost half of the universities that participated in the study. It is complemented by nearly two-thirds of universities conducting external audits to ensure IT compliance.

- **Inter-university collaboration:** The indicator "IT infrastructures (systems or applications) are provided to other universities" has experienced a 20% growth, reaching almost half of the universities that participated in the study.

### 5.3. Consolidated indicators

Some of these indicators, which are already of high value and have little scope for growth, were removed from the catalogue. These are indicators that almost all universities meet and are considered saturated. Other consolidated indicators include:

- **IT teaching support services:** Universities use an average of 6.33 ICT support services for teaching, which is an average of 90.38% of the total number of services included in the catalogue (7 support services for teaching). The percentage of teachers using the institutional virtual teaching platform is slightly over 90% in the last three editions.

- **Technological equipment of the classrooms:** At least 83% of the classrooms have an Internet connection for students and a multimedia projector. It has gone from 80% in 2015 to 81% in 2016 and 83% in 2017. Although already starting from a high value, it continues to grow slowly but steadily.

- **Institutional Repository:** The percentage of universities with an institutional content repository is 88% in 2017. In 2016 it was 82%.

- **Own and differentiated budget:** This indicator has been consolidated, standing above 90%.

- **Monitoring of projects:** The indicator "Do you draw up monitoring reports and at the end of the project your success or deviation from the initial objectives is evaluated?" has reached a stable value over time of more than 90%.

- **Performance of services:** The indicator "Reporting on the performance of systems and services in operation" has reached a stable value over time of more than 90%.

- **Report to the university management:** The indicator "IT area management provides the university management with up-to-date information on the status of IT" has reached a stable value in the time around one hundred per cent.

- **Collaboration between universities:** The indicator "Benchmarking in relation to other universities and assimilating their good IT practices" has reached a stable value over time of more than 90%.

The aim of the UNIVERSITIC report is to carry out a comprehensive and thorough analysis of IT in the Spanish university system from all possible points of view. In addition to its informative value, as a catalogue of indicators that offer a fixed picture of IT in the Spanish university, it also has a prospective character. In other words, it seeks to identify where universities have the greatest opportunities. For this reason, when designing a future dashboard to help identify opportunities, we are particularly interested in indicators with positive evolution and, within these, those that have the greatest potential for growth. Therefore, without losing sight of the consolidated indicators, the candidates for the set of KPIs in the dashboard would be the emerging and hot indicators.
6. CONCLUSIONS

Dashboards have become a very powerful tool for top management in organizations. In most cases, in order to build a really useful dashboard, a very deep analysis of the organization's characteristics and an iterative design process is necessary to lead to a totally customized tool. The ad-hoc design and deployment of dashboards is expensive and does not always end successfully.

In this paper we propose a classification system of indicators that are candidates for being part of the dashboard based on the values that indicators take over time and their trends. This is what we have called data-driven KPI classification and selection. This proposal is based on two stages: the classification of indicators according to their values and the selection of KPIs for inclusion in the dashboard.

For the classification of indicators, three categories are proposed, which classify indicators with a positive trend: emerging, hot and consolidated. This classification is due to the strategies of Spanish universities, in which we are looking for indicators that inform us of the variables with more positive evolution, although other categories can be proposed for indicators with negative evolution.

The selection of the KPIs that will eventually be part of the dashboard is made according to the organisation's strategies. In the case of Spanish universities, the main emerging and hot indicators have been chosen, considering that the consolidated ones have reached their full development and no longer have room for improvement.

The classification and selection proposal is a first step to automate some of the stages of dashboard design. In this way, in the future, dashboards will be able to be more dynamic, incorporating new emerging indicators and eliminating consolidated indicators that no longer provide information. In the future, we intend to expand the classification with negative trend indicators, to give a broader spectrum of indicators. We also aim to further automate the generation of dashboards.

7. REFERENCES


8. AUTHORS’ BIOGRAPHIES

Rafael Molina-Carmona received his B.Sc. and M.Sc. in Computer Science from the Polytechnic University of Valencia, Spain in 1994, and his Ph.D. in Computer Science from the University of Alicante, Spain in 2002. He is a professor at the University of Alicante, and he belongs to the department of Computer Science and Artificial Intelligence. He is also a member of the Cátedra Santander-UA de Transformación Digital, devoted to explore new trends in digital transformation. His interests are mainly the applications of Artificial Intelligence to different fields: computer-aided design and manufacture, computer graphics, learning, creativity, information representation and IT governance.

Faraón Llorens-Largo is a professor of Computer Science and Artificial Intelligence at the University of Alicante (UA), Spain. He received his B.Sc. in Education from the UA, his B.Sc. and M.Sc. in Computer Science from the Polytechnic University of Valencia, and his Ph.D. in Computer Science from the UA. He is the director of the Cátedra Santander-UA de Transformación Digital. He has been Vice-rector of Technology and Educational Innovation (2005-2012) in the UA, and Executive Secretary of the ICT Sectorial Commission of the Conference of Rectors of Spanish Universities (2010-2012). He received the “Sapiens 2008 Professional Award”, by the Valencian Official Association of Computer Engineers, and the “AENUI 2013 Award for Quality in Teaching Innovation”, by the Association of University Teachers in Computer Science. His works are framed in the fields of artificial intelligence, videogame development, the application of digital technologies to education and IT governance. Member of the GTI4U research team. More information in http://blogs.ua.es/faraonllorens.

Antonio Fernández-Martínez is a professor of Computer Science and Artificial Intelligence at the University of Almeria (UAL), Spain. He received his B.Sc. and M.Sc. in Computer Science from the University of Granada, and his Ph.D. in Computer Science from the UAL. He was director of the IT Service at UAL (1999-2007). He is currently the Government Coordinator and Delegate of the Rector for Interaction with Society and Companies of the University of Almeria. He is coordinator of the GTI4U research team, responsible for the research part of the UNIVERSITIC report for Spanish and Latin American universities and the IT Government Start-up Project, which has been successfully implemented in 10 Spanish universities. Both initiatives promoted by the ICT Sectorial Commission of the Conference of Rectors of Spanish Universities. Member of the BencHEIT initiative of European University Information Systems, of the ISO 20000 and ISO 38500 Standards Committee of AENOR and he is ISACA Academic Advocate.