DASH background

A central idea of DASH is that it should help users going from data to decision. To make this idea a bit more precise, in this document we explore the nature of decision making: on which levels are decisions taken, and what are the typical IT systems that support decision making.

In addition, we will try to position DASH in the spectrum of other Akvo tools, and to propose the top-level functional components.

Decision making and information needs

Decisions are made at all levels in an organisation, from the factory worker to the CEO, from the civil servant to the president. At these different levels, different types of information are needed. We can analyse the levels of decision making and the information required into three levels: strategic, tactical and operational.

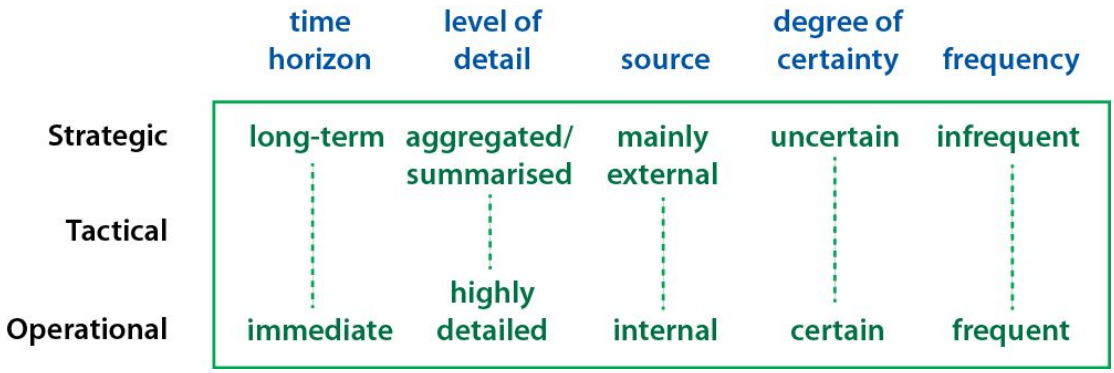
Strategic — On the strategic level, information over a longer timeframe from both internal and external sources is needed in order to plan long-term strategies of the organisation. This is about long-range decisions and planning of future projects. Internal information, both qualitative and quantitative, is needed in summarized form, in an easy to understand and appropriate format.

Tactical — The tactical level deals with decisions on the short term. It requires information from both the past and the present. It is mainly concerned with internal information, with only limited external information. Information is needed in summarized form, but detailed enough to allow tactical planning of resources and manpower.

Operational — The operational level deals with the performance of day-to-day tasks or an organisation. Information focusses on specific topics, and is obtained from internal sources. Information must be detailed, present, and precise, and is used for immediate operational decisions.

Obviously, the purpose of the information systems used by an organisation is to meet the information needs, to support the decision making process on all levels.

The chart below outlines the characteristics of information on each of these levels.



Typical examples of each of these levels include:

	Government (say, a Ministry of Water)	Business	NGO
Strategic	• country-wide service • coverage statistics • expected economic growth • private sector developments • donor requirements	• competitor analysis • overall profitability • future market prospects • total cash needs • expected government policy	
Tactical	• monthly forecasts • short term purchasing • requirements • project progress overview	• budget control reports • cash flow forecast • short term purchasing requirements • productivity measurements	
Operational	• purchase orders • asset lists • invoices • real-time functionality • work orders for repairs	• employee hours worked • inventory levels • daily sales results • reject rate	

A classification of information systems

One way to classify IT systems is to make a distinction between those systems that support the actual operation of an organisation, government or enterprise —*operational processing* — and those that track, report and analyze it after-the-fact, *analytical processing*. The division is slightly arbitrary as it is not always clear what is operational and what is analytical, but applications supporting each are called, respectively, operational applications and analytical applications.

However, it is possible to add a bit detail to this classification, based on the function the systems have. Below, we have listed six them in order of increasing complexity.

Operational systems

- **Office information systems** — supports regular office functions: word processing, email, calendar functions, spreadsheets, etc.

- **Transaction processing system** (TPS)— Transaction processing systems automate the handling of transactions, which are individual simple events in the life of an organization. These systems represent the lowest level in an organisation's use of information systems. They are used for routine tasks in which data items or transactions must be processed so that operations can continue.

Analytical systems

- **Management information system** (MIS) **—** A management information system extracts relatively raw transaction data from a database and convert them into a summarized and aggregated form for managers, usually in a report format. Used to make *routine* decisions. MIS systems are used to handle structured decision problems, where it is clear what the inputs and outputs needs to be.

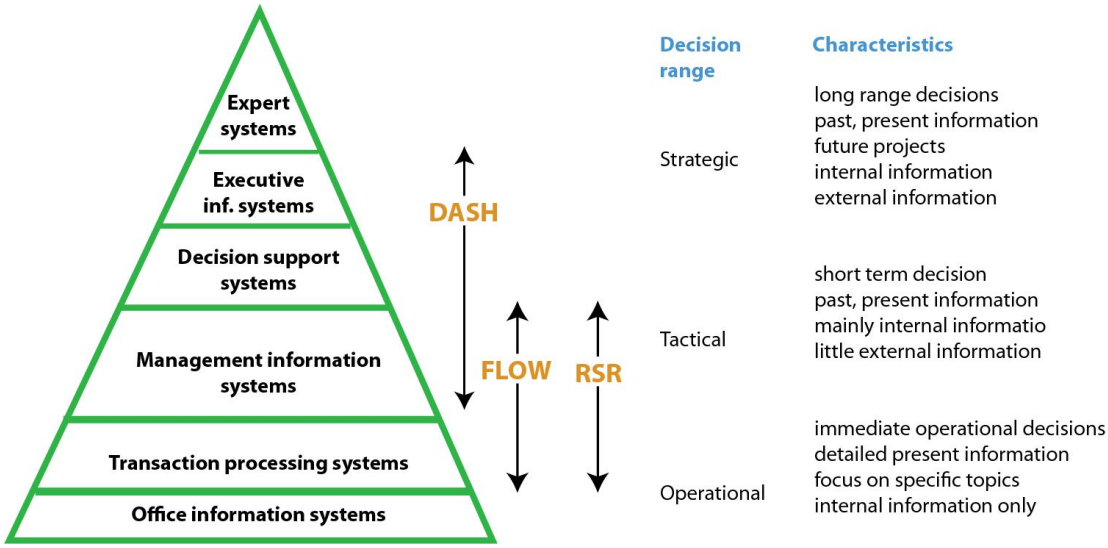
- **Decision support system** (DSS) — A decision support system is an interactive system that collects and integrates data from multiple sources to assist in making *non-routine* decisions. DSS systems are used to handle ill-structured decision problems, through direct interaction with data and analysis models. It must support interaction with non-technical users, should have access to a wide variety of data, and provide a variety of modelling and analysis tools.

- **Executive information systems** (EIS) — Provide managers with flexible access to information at the tactical and strategic levels. To develop plans and make major decisions, managers need relevant, timely, easily understood information, about both the internal and external environment of the organisation. In outcomes, they often resemble a Decision Support System, but where a DDS needs significant expertise to use, an Executive Information System is designed to help managers find information they need whenever they need it, in the most appropriate format. Management dashboards fall under this category.

- **Expert systems** (ES) — An expert system mimics expert judgment by following sets of rules that experts would follow; it relies on artificial intelligence — the science of developing computer systems that can mimic human behavior. Expert systems holds specialist knowledge about a specific subject area, and make it possible for non-experts to obtain information and recommended decisions from the system. Examples include systems found in law, taxation, granting credit in banks, diagnostics in medicine, and aircraft recognition in defence.

Positioning DASH

If we place the current Akvo products on this classification, the following picture emerges:



- **FLOW** spans the range between transaction processing system and management information system. It is mainly used to collect detailed information on a disaggregated level, which is then exported in Excel for further analysis.

- **RSR** also spans the range between transaction processing and management information systems. It is both used to capture detailed, disaggregated data on projects and project results, and, once the paperless reporting feature is implemented, can be used to create full reports on projects. However, it doesn't qualify as a decision support system, as those are more interactive and flexible.

- As foreseen at this moment, **DASH** spans the range from management information systems to executive information systems, as it can both be used to create simple routine reports, and be used to transform, merge and visualise data, and create dashboards for executive use.

An important fact is that DASH will not be used to edit or change data, but just analyse it — filter, group, aggregate, compute additional columns, etc. It will only operate on read-only data. Data cleaning, for example, should happen in FLOW, and editing project information happens in RSR.

Top level functional components of DASH

DASH is a data-analysis tool. Therefore, we can identify the following four main components:

- **Connect** — Connect to external data sources. This could be Akvo data sources, such as FLOW or RSR, or external ones, such as IATI, data.worldbank.org, NASA, weather data, population density data, etc.

- **Transform** — Filter and aggregate data, data facets, join data sets together, enrich datasets with geocoding and other computed columns (scoring, combined indicators, logic, etc).

- **Visualise** — Create dashboards consisting of different visualisations of data, including graphs, tables, charts, maps, etc.

- **Publish** — Make the visualisations and dashboards available for dissemination as website, iframe, widget, etc, on a variety of devices, and with a variety of access possibilities.

Characterising DASH

To succeed as a data to decision tool, we can formulate some high-level requirements[1]:

Visual and interactive — Interactive, visual analytics offers advantages to people of all levels of skill and training. With interactive visualization, it is possible to scan and analyze great volumes of data and to navigate through the data. Although a report can never scale to millions of pieces of information, visualisations can place this amount of data on a single page. Good visualisations allow the user to visually navigate and discover patterns in the data and relationships between data, and draw inferences instantly. This makes analytics an active process of discovery.

No coding requirement — To use traditional business intelligence tools often required substantial technical and programming knowledge, for example to perform database queries. The art of programming requires a great deal of skill and experience to be effective, and software artefacts need careful management to obtain quality and maintainability. The users of DASH, on the other hand, will usually not have such skills. To be truly useful, DASH should provide the full range of capabilities without the need for programming, scripting, or any other kind of code development. In addition, power users should be able to use coding if needed.

Conceptually sound —It will be very important to provide a clear conceptual model for working with the data. When you're not sure what the effect is when you click something, or if you don't understand the effect of loosening a constraint and sliding a slide bar, it's hard to use. The connection between the interface and the underlying conceptual model has to be obvious. Applying filters, constraints or other rules should perform accurately and consistently so the analyst can be confident in the answer and there is no need to double-check every result.

Because most analytical inquiry involves a series of steps, this conceptual soundness must apply across all the steps in the current analysis. This will be

---

[1] Inspired by: *The ten principles of enterprise analytics*, Neil Raden, http://bit.ly/tenprinciplesofanalytics

especially true for filtering, transforming, aggregating, and joining data sets. Conceptual soundness contributes to decision confidence. Likewise, there should be no need to understand (relational) data models. All references to the source or storage modes of the data should be abstracted.

Fast — Fast is always a relative term, but you know it when you see it. Interactive visualization and data analysis require near-instantaneous response time to keep up with the analyst.

Disposable/Persistent — It must be possible to conceive of and execute analysis quickly, and to dispose of the analysis when it is no longer needed. For those scenarios with a longer lifespan, a smooth mechanism must be present so that the steps in an analysis can persist and be reused as needed.

Collaborative — DASH should also foster a learning and collaborative environment. In the same way a reporting tool provides a means to distribute reports, an analytics tool should have the means to store and distribute analyses, a thread of reasoning that can be replayed and evaluated by others.

Safe and Secure — It should always be clear who has access to what data, and people should be able to trust that their data is safe. This means honoring and using standards for security and backups.

Expressive — Most importantly of all, for analytics to be useful, people have to be able to express themselves – the selection and managing of data, the design of relationships, calculations, hierarchies, etc., and the appearance of the visualizations. Moving from idea to visualization to action has to be simple and straightforward and the range of ideas and approaches should not be unreasonably constrained by the technology.

Sources of inspiration

Some products that have solved parts of the user interface experience we are after:

- Trifacta - for data wrangling, filtering, transforming, computed columns
- Google refine - for data filtering, data facets
- Klipfolio, Geckoboard, Kibana - for dashboards
- Tableau - for graphs and charts
- CartoDB - for maps
- DevResults - for maps and indicators