Association for Information Systems
AIS Electronic Library (AISeL)

AMCIS 1999 Proceedings

Americas Conference on Information Systems (AMCIS)

December 1999

Software Process Asset Representation Schemes,

Richard Gibson
American University

Software Process Asset Representation Schemes

Rick Gibson, Computer Science and Information Systems Department,
American University, Washington, DC, rgibson@american.edu

Achievement of Maturity Level 3 of the software CMM® requires development and maintenance of a usable set of software process assets. These assets consist of the following: the organization's standard software process, guidelines for the project's tailoring of the organization's process, descriptions of approved software life cycles, software process database, and a library of software process-related documentation. Zahran (1997) describes these components as follows:

- The organization's standard software process covers definitions and descriptions of the software process architecture and elements.
- Tailoring guidelines specify criteria for the tailoring needed to make the standard process useful and applicable to a specific project.
- Approved life cycle descriptions provide a reference or pointer to detailed technical documentation and available training for the life cycles.
- The process database holds process definitions and process performance measurements. The measurement portion should contain a few indicative, useful measures of the products and processes.
- The library of software process-related documents is where all process-related documents from completed software projects are stored.

To maximize usability, it is imperative that we be able to select among alternative presentation media for these various assets and, when graphics are appropriate, adopt the design principles needed to achieve graphical excellence, which is characterized by Edward Tufte (1994) as:"…complex ideas communicated with clarity, precision, and efficiency."

Media Choices

The Software Productivity Consortium offers a conceptual framework (Figure 1) for process representation with three distinct models--text, graphic and numeric. Contrast this with Tufte's (1983) contention that graphics "are paragraphs about data and should be treated as such." Words, pictures, and numbers should be integrated together, not separated into word processor documents, spreadsheet tables, and pictures.

Marks and Reiter (1990) suggest the following minimal subset of factors be considered in decisions regarding the choice between text or graphics for information display: constraints, type of information being communicated, and the expected use of the information.

Constraints may be imposed by the delivery medium or by the user population. If the information is delivered via a slow network link then text probably will be used. On the other hand, if the information were being delivered to an international user population that does not share a common language then graphics would be the preferred medium. A general finding seems to be that graphical presentations of information are often better suited to domain experts than to novices. All speakers of a given language, regardless of their domain expertise, possess a shared vocabulary of tens of thousands of words, plus a very rich set of syntactic, semantic, and pragmatic rules for combining these words into sentences. On the other hand, most people do not possess a similarly rich knowledge of graphics. This means that a text-based information-presentation system for people who are not domain experts can build on a rich existing knowledge of language, while a graphics-based presentation system must explain everything from scratch.

The type of information being communicated is a topic that has been discussed in the research community, although often from the perspective of individual systems. For example, Feiner and McKeown (1990) discuss media selection in a system that produces instructions for maintaining and repairing equipment. This system communicates location information and physical attributes with graphics, simple actions with a combination of text and graphics, and abstract actions purely with text. From a more theoretical perspective, Stenning and Oberlander (1995) suggest that graphical presentations do not always differentiate between what information is intentionally being communicated and what information the user should ignore. Physical information (such as location) is often best communicated graphically, while abstract information (such as causality) is best communicated with text. Graphics is better suited to conveying information about individual objects than to communicating information about classes of objects.

A final factor is the communicative purpose of the information (see Roth and Hefley, 1993). Graphics appear well suited to analytical tasks, because graphical presentations allow human users to exploit their visual pattern-recognition abilities. Similarly, graphics are well suited to marketing tasks, where a primary communicative goal is simply to grab and keep attention.

Textual presentations (including tables as well as natural-language text), on the other hand, are better suited for communicated information precisely, and for instructional contexts where a user is expected to memorize the information

The Software Productivity Consortium offers a conceptual framework (Figure 1) for process representation with three tracks--characterized by associated process improvement goals. Each track has six tiers, which establish an order of precedence for process representation development.

Figure 1 Goal Pairs	Disseminate and Train Process	Analyze and Improve Process	Plan and Control Process
Model Type Tier	TEXT	GRAPHIC	NUMERIC
1	Policies	Communication	Plans
2	Standards	Architecture	Schedules
3	Procedures	Interface	Cost
4	Methods	Flow	Status
5	Guidelines	State	Coordination
6	Instances	Simulation	Enactment

Principles of Graphical Excellence

In his three books *The Visual Display of Quantitative Information* (1983), *Envisioning Information* (1990), and *Visual Explanations* (1997) Tufte, a designer, statistician and Yale professor, suggests that bad graphics lie by distortion, obfuscate by omission and confuse by decoration. One chapter in *Visual Explanations* provides terse, stunning evidence of the real-world consequences of bad graphics display. It contrasts a Victorian doctor's successful use of maps to isolate the cause of a cholera epidemic with the misconceived charts used in the disastrous decision to launch the Challenger space shuttle. As Tufte (1997) edits and revamps the charts to highlight the cautionary logic in the numbers, you realize, if you hadn't already, that this stuff really matters. In analysis of the Challenger accident, it almost seemed as though the problem was engineers trying to present information design. Often people who see very well aren't very good quantitatively, and the quantitative people aren't very good visually. So, Tufte prescribes the following:

- focus on content
- compare rather than describe
- maintain data integrity
- maximize resolution
- use proven design concepts

A focus on content demands that above all else the representation shows the data. The focus should be on the content of the data, not the technique. This leads to design transparency. Avoid "fooling around with data" and use a clear, simple, straightforward design with a richness of data. The success of a graphic display is based on deep knowledge and care about the substance, and the quality, relevance and integrity of the content. Assume that the viewer is just as smart as you and cares just as much. Never `dumb-down' a graphic. Show the data in its full complexity and let viewers make their own discoveries.

Denouncing the presumption that graphics are "devices for showing the obvious to the ignorant," Tufte advises designers instead to assume a smart, curious reader who is willing to explore complex information as long as it is presented with grace and clarity. When viewing a visualization jammed with incomprehensible, cluttered graphics, there is a great temptation to remove data; even relevant information. But "clutter and confusion are failures of design, not attributes of information." If a visualization is too cluttered, don't remove data, change the design. This suggests a most unconventional design strategy is revealed: *to clarify, add detail*. Credibility comes from detail and in many cases one can clarify a design by adding detail. "High-density designs also allow viewers to select, to narrate, to recast and personalize data for their own uses. ... Data-thin, forgetful displays move viewers toward ignorance and passivity, and at the same time diminish credibility.

At the heart of quantitative reasoning is a single question: "Compared to what?" Most graphics today are descriptive rather than comparative. This may be part of the reason why scientific graphics, even those about multivariate phenomenon, are dominated by the xy-plot. The xy-plot invites reasoning about causality in a way that even the most impressive isosurface does not. We should strive for relational, rather than merely descriptive displays. Avoid relying on the viewer's memory to make visual comparisons; a weak facility in most of us.

Misleading displays are so common that Tufte (1983) has some very specific suggestions on this topic. The representation of numbers, as physically measured on the surface of the graphic itself, should be directly proportional to the numerical quantities represented. Clear, detailed, and thorough labeling should be used to defeat graphical distortion and ambiguity. Write out explanations of the data on the graphic itself. Label important events in the data. Show data variation, not design variation. The number of information-carrying (variable) dimensions depicted should not exceed the number of dimensions in the data.

The human eye has the capability to register 150 million bits of data. By comparison, the resolution provided by the most common displays of information seem inadequate, or at least inefficient:

- non-fiction best seller page 40 thousand bits
- telephone book page 80 thousand bits
- reference book page 240 thousand bits
- personal computer monitor 8 million bits
- high end workstation monitor 24 million bits
- 35mm slide 25 million bits
- large topographic map 150 million bits

Compounding the resolution problems associated with the computer screen is the allocation of only a small portion of the display to data and the rest to icons and other computer administrative debris. Railing against this "Operating-system Imperialism," Tufte asserts that only 40 percent of the screen is devoted to content, with the rest consumed by 5,000-pixel icons. What is worse still is the fact that the icons need a name underneath them so you can tell what they mean.

Tufte has researched a number of classic information designs and general principles. Some of these are small multiples, time series, and micro/macro composition. A small multiple design consists of a single design repeated several times within the eyespan, each example showing a different value of the independent variable(s). Use of small multiples allows comparison to be enforced within the scope of the eyespan. The time-series plot is the most frequently used form of graphic design. One dimension, usually the horizontal, is time, and the graphics march along showing variation as time proceeds. Finding innovative ways to incorporate time-series into graphics is needed. Micro/macro composition refers to an approach where a graphic contains enormous detail, but an overall pattern emerges.

Further support for Tufte's integrated text-graphic-numeric representation scheme is found in Caputo's (1998) criticism of a typical process definition outline that contains (among other items):

- Purpose, scope, objectives and definitions
- Activities
- Flowcharts
- Interdependencies
- Metrics

The resulting process definitions lacked clarity, precision with regard to who did what and why, as evidenced by an overuse of passive voice. She suggests using tables, with column headings such as "From Whom", "Activity" and "To Whom for What" to clarify and organize process definitions.

Conclusion

For purposes of discussion, this paper has presented two contrasting perspectives on the representation of software process assets. One approach utilizes a specific representation for each of the shifting goals of a process improvement initiative. The contrasting approach is to accept the premise that graphical excellence is the primary objective. Regardless of the approach selected by an organization, careful attention to the way in which process assets are represented will increase their use by project managers as they develop their project's defined software process.

In closing, it may be worth considering one possible resolution to the dilemma of representation scheme selection that emerges from a reconsideration of a scheme that is already among the most widely used--the flowchart. This useful approach embodies many of the characteristics of graphical excellence and allows a process to be depicted using one of three types of flowcharts (linear, deployment, opportunity), with a macro, mezzo, or micro level of detail for each. The linear flowchart displays the sequence of steps, which may help to identify redundancies or rework. The deployment flowchart identifies the people or groups involved at each step, which clarifies relationships. The opportunity flowchart differentiates activities that add value from those that add only cost.

References available upon request.