CHAPTER 17

EVALUATING INFORMATION SYSTEMS

An evaluation of information systems is a process to collect performance information and assess various aspects of information systems in order to provide useful feedback for better systems. Accordingly, evaluation is an impact study in its nature as opposed to descriptive research. This chapter introduces a general framework for evaluating information systems and discusses methods to measure performance of information systems and e-government applications.

17.1 Types of Evaluation

Since information systems in a broad sense consist of hardware, software, telecommunication, data and information, rules and procedures, and people (organization), evaluation need to reflect diverse dimensions and stakeholders.

17.1.1 *Ex ante* versus *Ex post* Evaluation

An *ex ante* evaluation is to provide policy information before a project for information systems is determined, whereas an *ex post* evaluation is performed after the project implemented. The former estimates costs and benefits of information systems, while the latter can use actual or more accurate costs and benefits although intangible costs and benefits are estimated. An ex post evaluation is *summative* (as opposed to *formative*) because it examines output and/or outcome of information systems. Also it is likely that an evaluation is conducted during the implementation of information systems to know how successful the implementation is.

17.1.2 Quantitative versus Qualitative Evaluation

A *quantitative* evaluation produces numerical form of data and employs statistical, mathematical, and computational data analysis methods. By contrast, a *qualitative* evaluation involves non-numerical data obtained largely from narratives and storytelling (e.g., participant observation and in-depth interview).[1]

17.1.3 Large N versus Small N Study

A large N study involves a large number of observations and is often related to quantitative research. A case study is the typical example of a small N study where a couple of observations are concentrated.

17.1.4 Internal versus External Evaluation

Evaluators can be internal employees or outsiders (e.g., consulting firms). *Internal* evaluation tends to be less costly and time-consuming than *external* evaluation, which can bring professional knowledge and independent perspectives.

[1] It is possible that a case study produces quantitative data such as a historical trend of population in the area.

17.2 Evaluation Methods

A research can be classified into experiment, observational study (i.e., survey and field study), or secondary data study (Singleton and Straits 1999). The last two methods are more frequently employed in social science than the experiment. Depending on the number of observations, a study is grouped into either large N or small N study although the magic number N is not clearly defined. A large N study collects many observations and often employs statistical and mathematical methods to analyze the data collected. In general, an interview and case study tend to be a small N study.

17.2.1 Experiment

In an experiment, researchers apply treatments to subjects (treatment group) in a fully controlled experimental environment and then compare the result with that of control group. *Randomization* and *matching* are employed to rule out any factor other than the treatments. Randomization assigns subjects into a treatment or control group randomly (as opposed to "arbitrarily"), matching matches treated and controlled subjects who share similar characteristics except for the treatment. If researchers cannot fully control the experiment, the research is called a quasi-experiment.

An experiment can be made in field or laboratory. This type of research tends to be quantitative. An experiment is not frequently used in social science due to its ethical issues in the research area.

17.2.2 Observational Study

Researchers observe subjects to get data without applying treatments to the subjects. Most of evaluation of information systems falls in this category. There are various types of studies under this umbrella. A study can be *participant observation* if the researchers participate in what is observed; otherwise, the study is *non-participant observation*.

Field research collects data directly by observing the target without applying any treatment. You may observe drivers' behavior in front of Urasa Station and count the number of drivers who violate traffic signals during 10:00-12:00 every day. Scientists may observe wild birds in Hakkai-san or count the number of pine trees growing in a hector. You may summarize what you have observed in numbers or narrative (describe in sentence).

Case study (ethnography) is to describe or explain a unit (e.g., a person, group, policy, town, island, state, or country).

17.2.3 Survey Research

Surveys collect information using such methods as questionnaires (by means of face-to-face, mail, email, Internet—Web-based survey), telephoning, interview, and others. The fundamental reason of survey is that researchers do not know the issue and thus ask those who know the issue well. Questions (survey instrumentation) are arranged in advance and respondents' answers are measured in Likert-scale (e.g., 4-point, 6-point scales). Some questions play a role of manifest variables or observed variables that are aggregated somehow in order to measure an abstract concept, which is called a latent variable (or factor).

Confirmatory factor analysis and structural equation model are used to analyze the relationship between manifest variables and latent variables.

An *interview* can be made in a form of face-to-face, telephone, or video (e.g., google video chat or Skype) to obtain in-depth information from interviewees.

17.2.4 Secondary Data Analysis (Literature Based Study)

This method analyzes various forms of documents and data that were generated by others. Examples are content analysis and meta-analysis. You may get survey data from PEW Internet & American Life Project (http://www.pewinternet.org) or tax related data from a tax department.

17.3 Subjects of Evaluation

What should be evaluated?[2] DeLone and McLean (2003) suggest information systems success model that includes system quality, information quality, service quality, use of information systems, user satisfaction, and net benefits (pp. 24-26).

System Quality	Information Quality	Service Quality
- Adaptability	- Completeness	- Assurance
- Availability	- Easy of understanding	- Empathy
- Reliability	- Personalization	- Responsiveness
- Response time	- Relevance	
- Usability	- Security	
Use	**User Satisfaction**	**Net Benefits**
- Nature of use	- Repeat purchases	- Cost savings
- Navigation patterns	- Repeat visits	- Expanded markets
- Number of site visits	- User surveys	- Incremental additional sales
- Number of transactions executed		- Reduced search costs
		- Time saving

Source: DeLone and McLean (2003: 26)

Heeks (2006) suggests four groups of evaluation indicators: readiness (awareness, infrastructure, digital divide), availability (supply, maturity stage), uptake (demand, usage, use divide), and impact (efficiency, effectiveness, equity)

Alshawi and Alalwany (2009) propose constructs of evaluation criteria grouped into technical, economic, and social issues (p. 201). Technical criterion examines performance (efficiency of services and personalized information and services) and accessibility (efficient user interface and disability access and language translation). Economic criterion measures cost saving in money and time. Social criterion evaluates openness, trust (in the Internet and government organization), and perceived easy of use and usefulness.

Table 17.1 summarizes evaluation criteria arranged by six components of information systems. For instance, public goods such as trust, openness, accountability, effectiveness, and equity are about people and can be used in cost benefit analysis. In software, usability,

[2] Educational assessment consists of formative and summative evaluations. Formative evaluation assesses need process (implementation), while summative evaluation includes outcome evaluation, impact evaluation (long-term net benefit), and efficiency evaluation (cost-effectiveness).

accessibility, and reliability are used in cost benefit analysis and their indicators may be generated by the information systems.

Table 17.1 Evaluation Criteria of Information Systems

Components	Evaluation Criteria	Costs/Benefits	Methods
People	Services (public goods) Satisfaction Usage	Cost and time saving Public goods	CBA Survey System-generated
Rules & Procedures	Security/Privacy	Security/privacy	
Data & Information	Data/information quality	Security/privacy	
Telecommunication Software Hardware	System performance - Usability, accessibility - Reliability, speed	Development Acquirement Maintenance	Benchmark System-generated

17.4 Evaluation of Information Systems

This section discusses frequently used methods to evaluate information systems.

17.4.1 Cost Benefit Analysis

Cost benefit analysis (CBA) studies costs and benefits associated with a project (program). Return on investment (ROI) and studying productivity (production function) of information systems are similar in a sense that they deal with monetary values. Related concepts are payback period and breakeven point.

In a CBA of information systems, it is relatively easy to estimate costs but difficult to estimate benefits because most benefits (e.g., trust, transparency, responsiveness, security, and privacy) are not tangible. In practice, the total cost of information systems tend to be underestimated by costs for maintenance, updating, and training (mainly direct costs are considered), while the benefits tend to be overestimated to make the cost benefit ratio larger.

Table 17.2 A Traditional Format of Cost Benefit Analysis (Categorical Accounting)

	Alternative 1	Alternative 2	…	Alternative *n*
Benefits				
…				
Subtotal				
Costs				
…				
Subtotal				
Net Benefit				

There are many stakeholders in public information systems including general public, users (applicants and government employees), public managers, politicians, and vendors. The CBA need to reflect the government perspective in favor of general public as the account domain; otherwise, CBA will be biased toward particular stakeholders of interest.

The next important question is “Who pays? Who gains?” In the information systems, the costs and benefits tend to be disproportionably distributed across stakeholders; information systems are financed by general tax but a small group of citizens can benefit from the systems.

The Kaldor-Hicks tableau proposed by Kerry Krutilla (2004), compared to the typical categorical accounting approach, enables researchers to analyze the distributional impacts on stakeholders in a public project (compare Table 17.2 and 17.3).

Stakeholders are listed in the column that show who will get and lose from the project (policy). Transfer is to transfer money from a stakeholder to another.

Table 17.3 An Example of Kardor-Hicks CBA Tableau

	General Public	Application Users	Government	Net
Benefit:				
Public Goods	G_p			G_p
Efficiency		E_a	E_g	E_a+E_g
Transfer:				
User fee		-U	U	-
Financing	-F		F	-
Cost:				-D
Development			-D	
Maintenance			-M	-M
Security			-S	-S
Privacy				
Net	G_p-F	E_a-U	E_g+F-D-M-S	$G_a+E_a+E_g$ – (D+M+S)

Source: Park (2004)

17.4.2 Users' Satisfaction

How many users have used information systems? How much have they been satisfied with the online information and services provided by information systems? These questions measure usage rates and users' satisfaction of the information systems.

An exemplary case is Annual Users' Satisfaction Survey of the University Information Technology Services, Indiana University, which have conducted for the past two decades. The survey takes a large number of random samples stratified by status (undergraduate, graduate, staff, and faculty) and campus (including regional campuses) and report average satisfaction rates of individual IT services on the Web page http://www.indiana.edu/~uitssur/.

This survey lists various information technology services that can be provided in a university.

17.4.3 Performance Indicator and Balanced Scorecard

Balanced scorecard (BSC), suggested by Kaplan and Norton (1996, 2000), is a formal approach to assessing overall organizational performance and keeping track of the execution of activities with respect to the targets set by the leader. BSC is top-downed and control-oriented in its nature. It provides a succinct summary (indices) of key financial and non-financial measures that shows performance information of an organization in a single concise report. BSC has financial (How do we kook to shareholders?), customer (How do customers see us?), internal business process (What must we excel at?), and leaning & growth (How can we continue to improve and create value?) perspectives.

Some performance indicators are provided by the information systems. Examples include CPU time (or CPU use), amount of data transferred, and the number of clients who use a particular online service. Of course, these indicators are used in a balanced scorecard.

17.4.4 Benchmarking

Benchmarking compares the best practices of providing online information and services across organizations or countries. Rankings announced by a couple of benchmarking have been sensational to scholars and practitioners in government.

United Nations Public Administration Network (UNPAN, http://www.unpan.org/) conducts Global E-government Survey and announces rankings of e-government readiness and e-participation. Darrell West at Brown University (http://www.insidepolitics.org) conducted global (federal and state) e-government survey during 2001 and 2008. The survey considers online information, electronic services, privacy and security, disability access, foreign language access, advertisement and user fees, and public outreach when determining rankings. Benchmarking focuses on supply or service provision side rather than demand side.

17.4.5 Stage Model of E-government Development

Many e-government development models were suggested to measure the progress of e-movement. These stage or phase models implicitly assume linear development of online services without considering technology-task fit (Coursey and Norris, 2008). Accordingly, it is absurd to evaluate e-government using this scheme.

Baum and Di Maio (2000) propose the first e-government model with four phases, which serves as a prototype for other e-government stage models. The presence phase is the simplest level, where basic information is provided on government Web sites. Citizens on the interaction phase are able to conduct search, download documents, and email to government officials. Transaction moves forward to make available entire transactions online. On the most advanced transformation phase, services are integrated and the relationship between governments and citizens is reshaped. As a phase goes up from presence to transformation, cost, complexity, time, and legal protection requirement also increase.

The United Nations (UN) and American Society for Public Administration (ASPA) (Ronaghan 2001: 8-14) suggest five stages of e-government development on the basis of e-government content and services available. The first emerging presence provides official, but limited basic and static information, while the enhanced stage offers dynamic and specialized information. On the interactive stage, citizens are able to enjoy the sophisticated interactions such as searching database, posting comments, and downloading forms and applications. The transactional presence allows complete and secure transactions to pay for services and conduct financial transactions (e.g., fines and taxes). The final seamless stage provides fully integrated services in a "unified package." Forlano (2004) employ this model to analyze OECD countries. Similarly, Hiller and Bélanger's (2001: 14-16) five stages of e-government include information dissemination, two-way communication (interaction), transaction, integration (portals), and political participation. The last stage provides online registration and voting, calling for special attention to the privacy issue. Moon (2002) uses this model to analyze municipal government Web sites.

Layne and Lee's (2001) four-stage model of e-government development is based on two dimensions of complexity and integration. The catalogue stage is very similar to the presence and interaction phases of Baum and Di Maio (2000). The transaction stage, common across previous models, is followed by integration for "one-stop shopping." Vertical integration across different levels of government within similar functionality precedes horizontal

integration across different functions and services (p. 125). West (2004) suggests four stages on the road to e-government transformation. The first billboard stage like the presence phase of Baum and Di Maio (2000) provide static information without two-way communication. Partial-service-delivery combines interaction and transaction in other models. The one-stop government portal provides fully executable and integrated services. The final stage is interactive democracy with public outreach and accountability enhancing features, where e-government "move[s] beyond a service-delivery model to system-wide political transformation" (p.17).

Table17.4 Stages/Phases of Electronic Government Development

Baum/Di Maio (2000)	UNPAN/ASPA (2001)	Layne/Lee (2001)	Hiller/Belanger (2001)	West (2004)
Presence	Emerging Enhanced	Cataloguing	Information	Billboard
Interaction	Interactive		Two-way communication	Partial Service Delivery
Transaction	Transactional	Transaction	Transaction	Portal Stage
Transformation	Seamless/fully integrated	Vertical Integration Horizontal Integration	Integration Participation	Interactive Democracy

References

Alshawi, Sarmad, and Hamid Alalwany. 2009. E-government evaluation: Citizen's perspective in developing country. *Information Technology for Development* 15(3): 193-208.

Baum, Christopher, and Andrea Di Maio. 2000. *Gartner's four phases of e-government model*. Gartner Research.

Coursey, David, and Donald F. Norris. 2008. Models of e-government: Are they correct? An empirical assessment. *Public Administration Review* 68(3): 523-536.

D. M. West, Global e-government report, 2001-2008, http://www.insidepolitics.org/

DeLone, William H., and Ephraim R. McLean. 2003. The DeLone and McLean model of information systems success: A ten-year update. *Journal of Management Information Systems* 19 (4): 9-30.

Gupta, M.P., Debashish Jana. 2003. E-government evaluation: A framework and case study. *Government Information Quarterly* 20: 365-387.

Hiller, Janine S., and France Belanger. 2001. *Privacy strategies for electronic government*. Arlington, VA: Pricewaterhouse Coopers Endowment for the Business of Government.

Kaplan, Robert S., and David P. Norton. 1996. *The balanced scorecard: Translating strategy into action*. Cambridge, MA: Harvard Business School Press.

Kaplan, Robert S., and David P. Norton. 2000. *The strategy-focused organization: How balanced scorecard companies thrive in the new business environment*. Cambridge, MA: Harvard Business School Press.

Layne, Karen, and Jungwoo Lee. 2001. Developing fully functional e-government: A four stage model. *Government Information Quarterly* 18(2): 122-136.

United Nations. UN e-government survey, 2001, 2003, 2008, 2010, 2012. http://www.unpan.org

West, Darrell M. 2004. E-government and the transformation of service delivery and citizen attitudes. *Public Administration Review* 64(1): 15-27.

Singleton, Royce A., Jr., and Bruce C. Straits. 1999. *Approaches to social research*, 3th ed. New York: Oxford University Press.