The value of information – the quest for ROI in organization models and software investments

Gabriel Anderbjörk
Stockholm, Sweden
http://gabriel.anderbjork.se

October 2009

Keywords: Software ROI, Information Economics, Knowledge Management.

***Abstract**: This paper decribes a model for valuation of information in a corporation.The model can be used to evaluate the efficiency of both software investments and organisational changes from a perspective of information management.*

The economics of information

The economics of information is a very complex topic, specifically in the large majority of organizations that have difficulty attributing revenue to be directly related to, or derived from, information. Indirect revenues related to information can however be, and usually are, substantial, albeit much harder to identify. The risk is thus imminent that the focus of corporate information economics is entirely the cost side of the equation. As there are no obvious revenues to match these costs against, decisions are taken to reduce, or even avoid, any activities that may increase the cost of information, decisions that sometimes can be fatal for the company.

Due to these problems, information economics must be separated into three distinct segments:

a) Directly measurable costs information.
b) Directly measurable effects of information.
c) Assessed effects of information.

Directly measurable costs of information

The direct costs of information can be separated into four categories:

- Purchasing or production costs
- Storage and access costs
- Quality costs
- Security costs

Purchasing or production costs are direct costs related to the acquisition of information. When purchasing information this is apparent as the cost equals the price invoiced by a third party. When produced internally the issue of costing is split. Information is produced for a primary purpose as well as byproducts of other processes. The cost for information produced for its own sake is the full cost of time spent, tools used etc. The cost for information that is produced as a result of other processes is much less specific. The norm should be that unless the information produced is the primary cause for the production, it should be considered entirely as a byproduct. The consequence of this is that the assumed cost of the information should only be any marginal cost (if any) associated with the capturing or storage of the information, and certainly no part what so ever of the cost of the main process from where it is derived.

Storage and access costs are direct derivatives of costs for infrastructure (hardware, software maintenance staff and monitoring systems) and physical storage rooms and storage media such as tapes and back-up discs. To measure such costs are usually quite easy as they are represented by defined cost units in an organization and it is unusual that these units have other side mandates that significantly impacts the cost level. The challenge is to allocate these costs between different information resources and, in worst case, different information types. It is not unusual that information resources with limited usage and their support systems are allocated a disproportionate amount of the total cost base. This can be alarming as the fact that they are limited in usage sometimes also means that they account for *relatively* low "revenue values". In such cases, wrong decisions might follow unless the cost reflects a real use of resource. To argue for a one-model-suits-all regarding allocation models on an aggregated level is quite obviously not possible. Such allocation must be made on an operational level when the real resource requirements are known.

Quality costs are often referred to as costs sprung out of failures or malfunctions as well as ongoing quality checks which are important to monitor data is of high quality for reliable decision making. In this case we do however refer to such costs that are derived from information lifecycle activities aimed at ensuring a high level of information quality. Quality parameters of information differ from organization to organization but typically include:

- Relevance
- Level of accuracy
- How current the information is
- Accessibility

Costs associated with ensuring these parameters are typically related to personnel, such as labor costs (hourly rates) as well as access costs which are typically system and terminal costs.

Security costs include costs related to limitation of access (e.g. systems for IDC – Intrusion Detection Control), encryption or, in advanced applications, electronic counter measures (ECM) or even stealth costs.

Directly measurable effects of information

It is indeed very tempting, in contrast to information costs, to talk about information revenues. If however one does not belong to a media company, which charge for information as a commercial service or product, it would be a misleading terminology. It is therefore suggested to refer to these as *information effects* which include effects (by default assuming positive impacting effects) on processes, revenues and even costs.

Factors determining the monetary value of a specific information element are usually defined as purchasing costs, storage costs etc (see above). These costs do not derive the *value* of the information as this is determined by the consumer/end user of the information and not at the point of production.

There are three prerequisites that are key to the value creation of any information element. These are:

- Timing

Did the information consumer get access to the information at a time when it could still impact a decision? If not, the information element was entirely without value *at that consumption point and at that time*.

- Relevance

Is this information relevant for the consumer's decisions? If not, the value could be negative as it may be considered noise or disturbance.

- Competence of the receiver

This third prerequisite has no real baring on information management itself but is still a very important factor to consider as a relevant information element may reach the right person (from an organization perspective) at the right time but if this person happens to be unable to utilize the information for a variety of reasons, e.g. the person is new to the organization or because the person fails to see the relevance in the information element, no decision will be affected and the information passes by without any value.

So, how then do we measure the effect or value of an information element? Provided that relevant information was received in time, was of good quality and reached a suitable recipient. The value is entirely dependent on the effects on decisions the information has given.

As an example:
If we can claim that information element X was the unique difference that enabled a sales rep to win a contract in front of a competitor, the value of that information, *at that location and at that point in time,* is exactly equal to the gross margin of that won contract. With the information – contract, without the information – no contract. Now, the world is never that black or white but this line of thought provides a clear model for how to approach the exercise of valuing information. Further, the expression " *at that location and at that point in time"* is very important to recognize as each information element can be consumed over and over again at different locations and at different times and it is this potential multiple effect that makes economic valuation of information both complex and important.

The unfortunate reality is that the number of directly measurable information effects is limited provided that a measure of the full effect is required. However, it is quite possible to assign a minimum measurable value on certain information elements. The prerequisite for success is that the information element's level of effect is rather obvious. Examples of such effects are time measures (information resulting in time savings) or cost savings (clear effect on lowered purchasing price or similar). The paradox is that such information is usually very specialized and hence does not offer a high probability of multiple effects. Another difficulty measuring the effects of this kind of information is that it usually corresponds only to a very small part of the total potential effects of information. In absolute terms it might equate to significant numbers but the big gains are more often to be found in the assessed effects of information.

Assessed effects of information

The very same prerequisites are applicable to the assessed effects of information as for the measurable effects, i.e. time, relevance and competence of the receiver. If these three are fulfilled, it is highly likely that the information element contributes with a positive value. However, to work with assessed effects on the same level of detail as one can do with measured effects is practically impossible. Hence, for such information, the model has to consider more aggregated levels such as information types or even full information resources.

The information equation

Valuation of information is still an area with lots of room for further research and it is also very much an area within which more or less serious debates take place as it is far from an exact science. During the Nineties comprehensive models for valuation of competence capital and structure capital in large corporations were discussed. Today a rather accepted view of these models tend to be that they were far too comprehensive to actually be of any practical use as the level of aggregation was too high to serve as strategic or tactical decision support. With polite criticism to our colleagues from the previous decade one could argue that the trap they fell into was to consider *information* as not enough value to address on its premature state but rather focus on the much more mature, complex and immeasurable resource labeled *knowledge*.

As a contrast, this paper puts forward a very concrete model expressed in the following equation which is called the *information equation*.

$$V_i = \sum_{X=1}^{n} \Delta R_X + \sum_{X=1}^{n} \Delta C_X$$

Vi: Monetary value of information element ”i”
ΔRx: Change in revenue at decision point x due to information element ”i”
ΔCx: Change in cost at decision point x due to information element ”i”

In order to be able to make any qualified assessments of information effects one has to work in multiple dimensions and also seek support in some simple probability theory. A schematic model like the equation above connects available information types to known decision areas in a matrix in which the matrix elements are assigned different weight values and probabilities with respect to their effect on the decision area at hand. Further, each decision area has a value for its unique *potential of effect* per time unit (e.g. annually) in a measurable unit (such as money or hours). A fundamental requirement in this model is that this potential of effect should be assigned as the estimated delta between current state and potential state. E.g. if a decision area covers sales of approximately 100 M Euro and it is estimated that the true potential (as an impact of the information) is 130 M Euro, the potential of effect equals 30 M Euro.

In the next section a matrix model is presented, extending and representing the logic described above in a more operational view. Still, it is important to emphasize that, the matrix language and mathematical logics aside, it is still a question of estimates and probabilities and not an exact science. When the matrix is filled it serves as a structured guidance regarding which information types that ought to be prioritized when it comes to purchasing information, systems for collecting, managing, disseminating information or any other infrastructure related investments or management costs. The model is supposed to serve as decision support for prioritizations and investment decisions and should be used for structuring and simulation exercises to built arguments upon, both for and against such decisions.

The information valuation matrix

When building this matrix, the approach should be to define those Information Types (IT_{1-n}) that are to be tested against the Decisions Areas (DA_{1-n}) within which they are assumed to have an impact. Within each DA the potential of effect (Δ Pot) is then estimated to its full value. Thereafter, the relative weight value (W) between the different Information Types is estimated per DA plus the probability (P) of impact for each Information Type in each DA. The sum of all W for each specific DA must be 1.0.

For simplicity in the matrix we label Δ Pot as X and also note that Δ Pot theoretically equals the total estimated ΔR+ΔC in the information equation for the entire DA.

		Decision Areas				
		DA 1	DA 2	DA 3	DA 4	DA 5
		Δ Pot: = X_1	Δ Pot = X_2	Δ Pot = X_3	Δ Pot = X_4	Δ Pot = X_5
Information types	IT 1	$P_{11} \cdot W_{11} \cdot X_1$				
	IT 2					
	IT 3					
	IT 4					
	IT 5					
	IT 6			$P_{6\,3} \cdot W_{63} \cdot X_3$		
	IT 7					
	IT 8					
	IT 9		$P_{92} \cdot W_{92} \cdot X_2$			
	IT 10					
	IT 11					
	IT 12					

Now, how to actually ***do*** this practically?
As already stated above, this is a structuring tool for assessing values and comparing consequences of potentially decided investments and operations activities. Hence, it is not a task for a one-man analysis. The process of building a matrix includes the “wisdom of the masses” and should be performed in teams. The starting point could be either to call into question the management of certain information types or the performance of certain decision areas; all depending on who it is that calls for the analysis. Just to illustrate the process, let’s assume we are a provider of equipment to a large industry segment (e.g. telecoms, utilities, pharmaceutical or alike). This would typically mean that we have a range of products offered on a range of geographical markets, most likely with some global competitors but also with a range of local competitors on each market. Our customers may purchase our products centrally or locally depending on the type of product and each customer’s purchasing policies. The product manager for product Q believes that, despite adequate marketing, we are underperforming in sales in five European countries. The assumption is based on both external market analysis and internal benchmarking of product features and evaluations vis-à-vis competition. We decide that one DA represents *sales of product Q on one market*. Hence, we end up with five DA’s for this valuation.

The next step is then to agree on the Information Types to be valued. Fundamental microeconomic theory states two lemmas;

- Actors have access to perfect information
- Actors take perfectly rational decisions and optimize their production

These have never been 100% true but we are getting closer to the first with all information that is available commercially as well as from open sources. The challenge is to define which information types that would be considered perfect for the decision area in question.

In our case, with an increased sales target on the table, it would be very convenient to just ask for our competitors pricing and believe that it would solve our problem but the sales process is of course far more complex. If our decision area is sales of a particular product on a particular market, the number of decisions to be taken during the process from leads identification to contract signing is significant, all with the same end target which is to submit a winning proposal. In this particular case, the first step would therefore be to send a very simple question to all sales individuals covering the related markets with the product in mind, namely “what hinders you from selling e.g. 30% more of product Q?” There are two typical answers to this question; a) the market is not there, or b) we are *lacking the proper sales support*. Now, as we have already decided that the market is there, this should be clearly stated in the question e-mail to disqualify that option. We will then be left with a range of answers in line with b above (no one will answer “I am not doing my job properly”) and that is exactly what we are looking for. A majority of the answers will prove to be information related; we know too little about…., we are lacking information on…, competition knew something we did not know…, etc. More concretely we are expecting requests such as more case studies and better understanding of the customers’ own market (to be able to be more persuasive that our product will contribute to the customer’s profit making). The next step for the person doing this analysis is to translate these answers into concrete information types that can be evaluated. The level of detail in this definition work is very much up to the analyst or the product manager in charge.

We now have the three key “variables” in the matrix;

- DA_{1-5}
- X_{1-5}
- IT_{1-5}

In this particular case though, keep in mind that X does not equal the total expected sales increase, it equals the total increase in gross margin that such sales increase will result in!

The next step in this process is split in two parallel tracks that are not to be mixed. Firstly, sales managers and possibly marketing managers for the DA markets are to suggest weights (W) for the different IT’s on their respective markets. At the same time, a new mail is to be sent to the sales staff asking them to set a probability number (P) (0-100, later to be divided by 100) measuring on average over the last 12 months, how likely is it that each of the IT’s would have contributed to more sales. For each P-value, take the average of the responded values and enter into the matrix together with the corresponding W’s and X’s.

At the end of this exercise we have a wealth of data in front of us. The obvious results are to sum up the rows for each IT and see the value each such IT is estimated to contribute and consequently ensure that those with a positive net contribution (estimated information costs deducted) are provided to the sales teams. Secondly, by comparing the average of P’s for each column we can identify potential differences between the geographical markets that are beyond the information matter. If one market is significantly lower than all the others, one might ask why? Are there other matters we have to address? Is it possibly so that that particular market, despite the original analysis, is less attractive that the other four? Comparing the allocation of W values from the different markets may also tell us something more than just the mathematics will do. Why is IT5 valued much higher on market 2 than on the others? Can we learn something from that?

The information valuation matrix is a format for the valuation of information in business processes. Information is a very valuable and tangible resource and the hope of the author is that this model can help in managing this otherwise often undervalued asset.

In the example above we used a situation with a lot of individuals involved. Another example is the CTO of a large European manufacturing company who estimated the potential of effect to be 200 M Euro annually only within the decision area process support to the design engineers. A less theoretic way to put the same statement would be to say that if the relevant, and asked for, information elements were available for each design engineer at each point in time of the design process, the annual net profit of the company would increase with 200 M Euro. Translate that to Return on Capital Employed, Profit per Share or p/e ratios and it is quickly realized that such information management will have very tangible effects on measures that keeps board of directors and shareholders happy.

However measured, there is no longer any doubt that information management as an executive managerial discipline hides substantial gains to be uncovered.