The Dynamics of Analytic Collaboration

Michael E. Fletcher

Abstract: *This paper examines the nature and effects of collaboration in a dynamic information environment using a System Dynamics Model. We model two analysts who are attempting to acquire knowledge. We ask how successful they are as a function of how much they already know, how skillful they are, how fast their document set is changing, and how much they collaborate. Our findings indicate a rationale for collaboration under conditions of high operational tempo. However, we also find reasons not to collaborate under other conditions. We point out connections between the rationales under our model and the behavior of real social systems.*

Introduction

Collaboration is currently a hot topic in government and industry. Organizations want to build and leverage collaboration in order to enhance their effectiveness. Collaboration is also of interests to professional analysts who must contend with information that is often doubtful and almost always perishable. Analysts want to use the high confidence knowledge of their collaborative partners to more rapidly build their own understanding.

Many sources discuss the virtues of collaboration, and often recommend strategies to encourage it. Unfortunately, none that I could find went into any depth on the dynamics of collaboration. The literature, which spans several disciplines to include psychology, business, knowledge management and organizational learning, is complex and often contradictory. I was unable to locate any reasonable discussion on how collaboration actually *works;* that is, provide it a meaningful operational definition. The model was therefore my own effort to fill a gap in my own understanding.

Summary of Definitions

Working with concepts that potentially span several fields can cause problems with definitions and frames of reference. To avoid this I would like to establish some working definitions.

Collaboration

Collaboration is defined as a group sharing resources and/or working together to reach some shared goal. Collaboration is a feedback process. Defining a simple System Dynamics structure to represent collaborative knowledge building was the primary purpose for the model.

Analyst Knowledge

Knowledge is the primary resource that needs to be shared in order to build analytic collaboration. There are many other types of resources that could be shared in a collaborative effort beyond knowledge, but knowledge is the only resource I intend to examine in any detail. In the language of System Dynamics, knowledge is a stock accumulation. It is also necessary to define, as simply as possible, the real or practical limits to knowledge within a dynamic information environment.

Learning

Learning is a rate of change in respect to knowledge. A simple model structure for analytic learning has to be defined. It is necessary to point out that many mental models regarding “knowledge transfer” are at worst faulty, and frequently vague. In my investigations I found that collaboration, “knowledge transfer,” and even Knowledge Management were often not provided meaningful operational definitions. Definitions or discussions frequently violated stock & flow physics.

Collaboration and “knowledge transfer” are not magical processes. What is in my head is not suddenly transferred to your head. Each individual must learn independently – my knowledge cannot increase your knowledge; however, my knowledge can increase your *rate of learning.* This is a key distinction, which is often poorly understood. I was surprised to find that establishing a relatively simple System Dynamics operational framework for collaboration largely demystified it and provided the best insights.

Relevance
Part of establishing an acceptable System Dynamics structure of the learning process is to establish the real or practical limits of learning. I wanted to examine collaboration in an environment where the information relevant to the problem at hand is rapidly changing. While modern information systems can deliver more raw data than any human can possibly consume, the amount of information directly relevant to the problem or question at hand is often severely limited. To represent this I used the concept of "average time to lose relevance" - essentially establishing a time constant for information of relevance. Admittedly, this is a simplification.[1] The model uses an aggregated time for relevance loss for all current knowledge. In reality, this aggregation is made up of subsets of knowledge where that figure might vary considerably.

Methodology

Purpose:
My purpose was to better understand how collaboration *works* by creating a System Dynamics model to examine its structure. Creating background of a dynamic information environment abstracted the difficulties of two analysts attempting to learn collaboratively. Not wanting attempt an epistemological definition of what "knowledge unit" might be, the units of measure used is generic. "Documents," in this case, could represent any quantifiable source that can be used as a basis to build knowledge.

Method:
Using a small System Dynamics Model I will investigate what can be learned about collaboration under different conditions.

The Model consists of three main parts:
1. Document (Information or Data) Flow.
2. Analysts attempting to gain knowledge about that set.
3. Collaborative interaction between two analysts.

The Implications of Model Boundary:
It is not the purpose of the model to completely represent the many complexities of collaboration, knowledge flow or learning. The purpose of the model is to capture just the key dynamic relationships, and explore the implications of how their interactions play out over time. This results in simplifying assumptions, which need to be made explicit.
1. We are assuming the analysts have shared goals and context for communication. These are normally required for collaboration.
2. The lack of shared goals and context are important problems in collaboration but are not subjects of this investigation.

Assumptions:
All types of analysis may contain unstated assumptions, often because the analytic boundary is not defined or poorly defined. These potential pitfalls make it useful to state ones assumptions up-front.

The Model is notational and doesn't pretend to represent any specific real world case. The rates of learning specified for "experienced" analysts are in many respects arbitrary, but are set to levels where that analyst will begin to fully understand the document set (read: analytic topic) within one to two years. That time frame is in line with my estimates on how long it takes an analyst to become familiar with a new topic.

In our model we start with two experienced analysts working a new document set, which covers a topic about which they initially know almost nothing. We then examine how collaboration affects different possible scenarios in respect to operational tempo, ability etc.

To a large extent the constants define the model boundary.

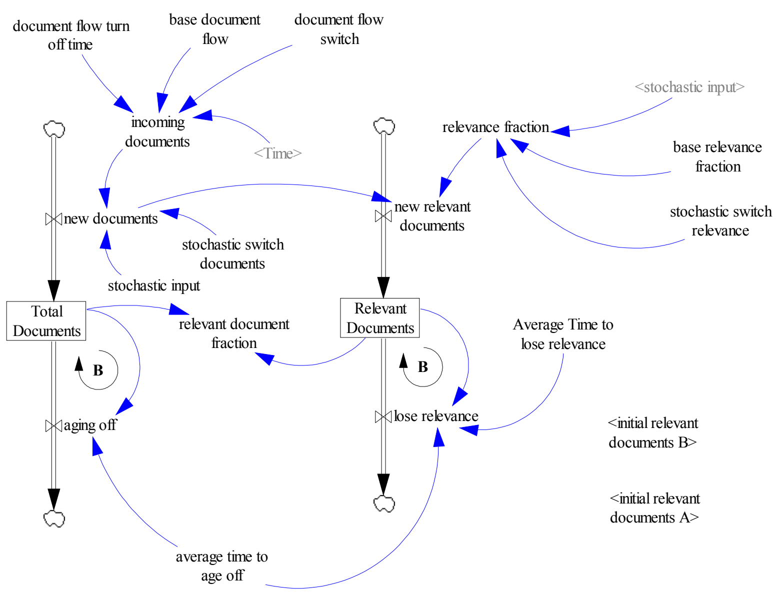
1. Time delays are for the most part constant.
2. There is only a small delay in the transference of information between analysts. This small delay factor (represented by a smooth) does not significantly alter the dynamics. This implies that even whatever knowledge is transferred is conserved; that is, it is not subject to the losing relevance[2] like the document stocks.
3. Relevant documents come in at a fixed rate. Changing input to some moderately random (stochastic) input does not alter the dynamics. The main reason for this is that the model stocks buffer their input rates due to the effects of integration. Additionally, the structure of the model is goal seeking; it does not have the structure to allow it to oscillate. It will continue to seek that goal whether the input is randomized or not. The underlying behavior modes of the structure are what are of interest, and randomizing the input only obscures the dynamics without changing the behavior modes resulting from the model structure. The overall dynamics weren't significantly affected by randomizing the flow rates, but there was one minor impact. Randomizing the inflow rates for documents or relevant documents did make the stocks slightly more resistant to going to zero because of the effects of integration. That is, the document stocks did not decay to zero as quickly when *average time to lose relevance* was low. I didn't consider this interesting side effect of the math to have major impact.
4. I did not impose a limit to how many documents can be converted per time period, because analysts rarely suffer an "embarrassment of riches" in such matters!
5. Values for *Query Effectiveness* and *Learning from Recovered Documents* are constants. In the real world these would likely improve gradually over time as the analyst gains experience. The time frame over which this happens is generally so long as that the impact on behavior is low. Also, improvement there would only enhance the results of collaboration and not degrade it.
6. There are only two analysts in this model – the multiple feedback effects of multiple knowledge resources feeding back to enhance learning is outside of the model boundary. (See section on follow-on work.)
7. The MIN function is used to ensure that total documents used to calculate "learning from recovered documents" does not exceed the total possible number of documents under extreme conditions. The function is again used to forestall the unlikely, but possible case that "age off time" is shorter than "time to lose relevance." In both cases extreme values are unlikely to be seen in real-world scenarios.
8. Except where specifically mentioned the Analysts are working with the minimum of initial knowledge. A value of one document is used as a starting point.

<u>The Dynamics of Document Flow</u>

In order to have a model of Analyst learning we must have information sources for them to access and attempt to learn from. In this model we use new documents flowing into some repository as an analogy for any information or data source. Some small fraction of those incoming documents is assumed to be relevant to the analyst's problem. Once those documents are recovered the Analyst has gained knowledge. In our model the flow rate is 1,000,000 documents/week and the relevance fraction is 1/10,000. The relevance fraction determines the relative "richness" of the document set. These are arbitrary figures for our notational model.

The Dynamics of Document Flow

(Figure 1.)

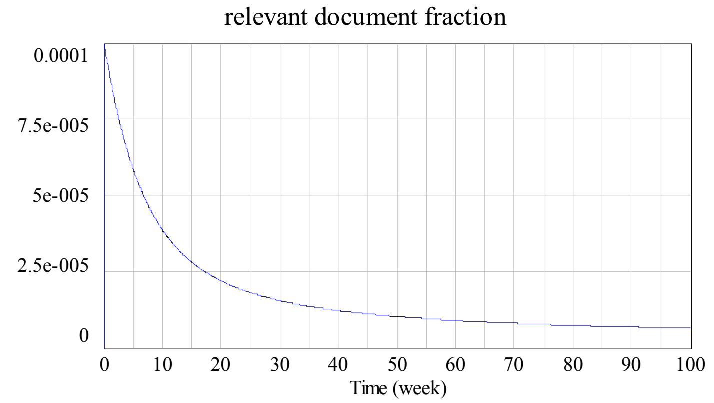


It is necessary to discuss briefly how *average time to lose* relevance affects the stocks. One of the misconceptions regarding information accumulations is that once information is placed in some big database, information becomes static. Nothing could be further from the truth. Because the relevance of documents tends to erode over time, at least as far as analysts are concerned, document repositories are dynamic creatures even if nothing is being added or deleted.

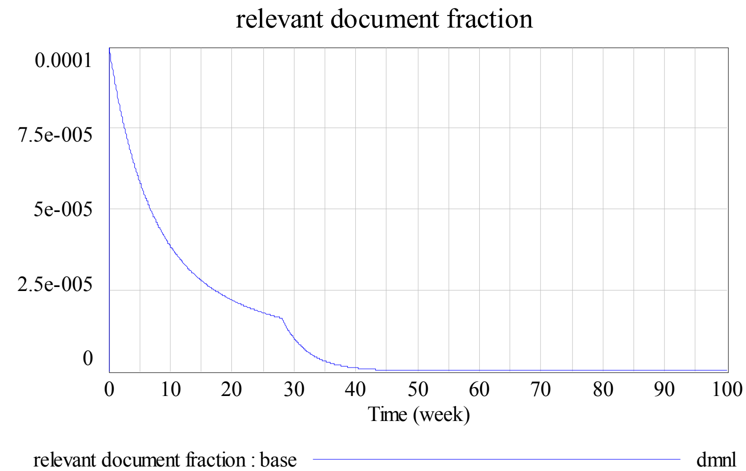
Even this very simple document flow model has some interesting dynamics. Because the values for *age-off time* and *average time to lose relevance* are different, the richness of the overall document set changes over time. Analysts intuitively know that "new" document sets tend to have more richness than "old" document sets. Additionally, analysts intuitively know to quickly drop the use of document sets, which no longer have new data going into them. This is because; *while the document set may be static in content it is always dynamic in relevance.*

If the relative richness of the document set is known and dynamic relevance of a document set can be measured (in *average time to lose relevance*) the relative worth of a document set can be calculated using simple calculus. If the inflow into a document set is turned off, the relevance of the documents will decay to essentially zero after three time periods (*average time to lose relevance*.)

The fact that document relevance generally decays at a different rate than the document set (age-off time) means that the dynamic document sets relative % of relevance usually starts high and then decays to some low level which is retained over time.



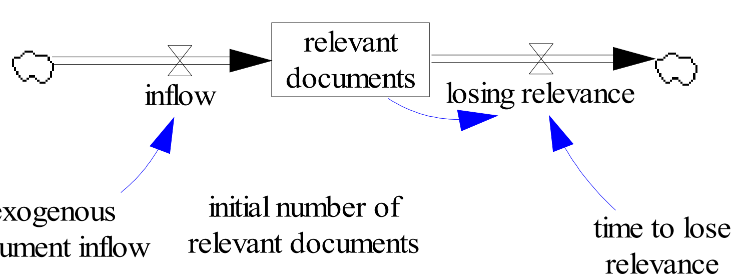
"Turning off" the inflows into a document set degrades it relevance rapidly – it will decay to essentially zero after three time periods. (average time to lose relevance)

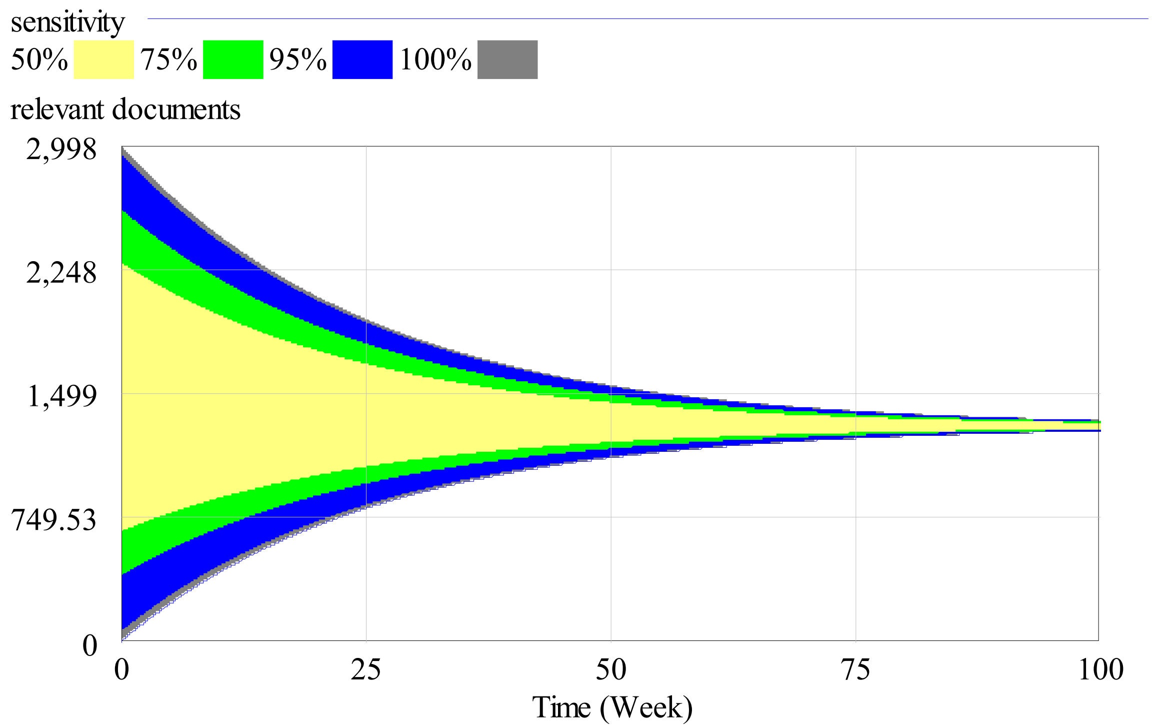


(Figures 3 & 4.)

It is generally understood that the dynamics of embedded simple generic structures can have a profound impact on even the most complex system. This model is no exception. One effect of dynamic document relevance is that even if the rate of incoming relevant documents is fixed, the *average time to lose relevance* dictates how many documents are available at any one time. This is because the *average time to lose relevance* determines how fast the stock of relevant documents is drained, and along with the document inflow, strongly determines the behavior of the relevant document stock. When an Analyst complains that operational tempo has eroded the relevant knowledge about a problem, that analyst is essentially correct. We can represent this effect with a separate simpler model shown in Figure 5.

(Figure 5.)





(Figure 6.)

50 docs/week inflow, varying initial number of relevant documents 0-3000, 26 week time to lose relevance

The sensitivity graph in Figure 6 shows that the *stock of relevant documents will seek the value of the inflow rate (documents/week)*time to lose relevance (week) without regard to the initial value of the stock.*

I will risk a short analytic comment to discuss Figure 6. I linger on this topic only in order to provide the reader context in regards to the thinking behind using the concept of "time to lose relevance." The reader might interpret the sensitivity graph to imply that all initial knowledge of an analytic problem decays over time, and therefore the initial knowledge of any problem loses its usefulness as time passes. This is only partially true. As discussed earlier, *average time to lose relevance* is a simplified aggregation. This mitigates against but doesn't completely falsify the implications of the dynamics in Figure 6.

In a larger sense, we are often strongly biased towards overestimating the value of our initial knowledge when approaching any problem. It is impossible to have a completely up to date understanding of the current state of any dynamic system. There will always be an information delay – no matter how small. We make up for this shortcoming by often using past data or history as a source of analogies to estimate the current state of the system and we use the same to extrapolate trends. This often suffices, (or perhaps satisfices), but is also the source of much potential error. Of course, things do tend to go along very much like they have in the past – until they don't. Writings in other fields tend to be mixed in regards to the value of initial information. The noted 20th Century Philosopher Karl Popper persuasively warned of the Poverty of Historicism (Popper, 1957). Political Scientists and Historians, who undoubtedly have a strong vested interest in maintaining the perceived value of history as a learning aid and decision making guide, provide only occasional guarded warnings. (Neustadt, 1986)

What we label as relevant to understanding system behavior varies. What an analyst evaluates as relevant to the understanding of system behavior is a key discriminator in determining overall analytic skill. It is prudent to define the boundaries of analysis, before examining the data, because the boundaries define what data will be examined to build learning. Or as Einstein said, "Our theories define what we measure."

Our knowledge of system behavior often concentrates on observed or observable behavior modes and "events" which have a relatively short persistence over time. Being an expert on current events does not define a person as an expert analyst! The limitations of our ability to maintain learning rates and knowledge levels in the face of an often rapidly changing information environment make relying exclusively on this type of analytic strategy extremely dubious.

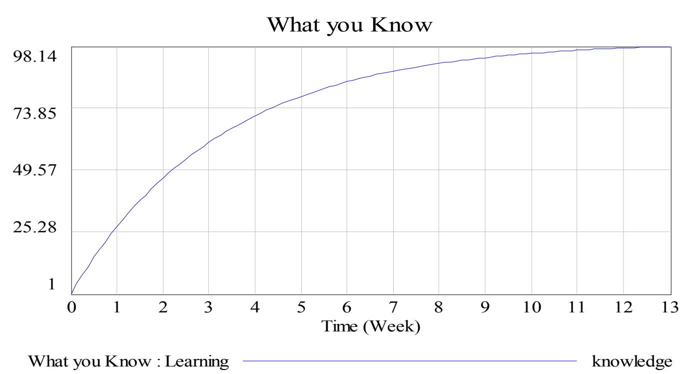
The key to being labeled an expert analyst (or a good system dynamics modeler) is to instead concentrate on the parts of the system, which have a relatively long persistence over time and retain relevance to system behavior over long time horizons. Using this strategy the "stock" of relevant knowledge that the expert retains is much more resistant to decay.

Learning Curves

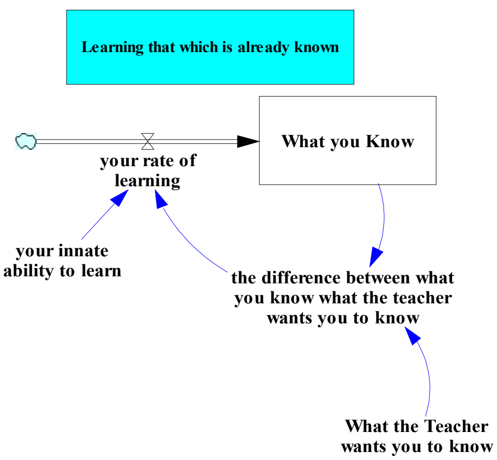
Even though the phrase "learning curve" is in popular usage, surprisingly, there isn't total agreement on what the learning curve of different types of learning looks like. An in depth discussion of the various types and dynamics of learning are beyond the scope of this paper; however, a sensible representation of Analyst Learning has to be developed in order to represent that process in the model.

Thankfully, there is general agreement that learning is not something that can be represented with a straightforward linear equation. There have been several attempts to approximate the shape of and define an equation for the learning curve in System Dynamics and related disciplines. Learning is known to be a feedback process. One of the earliest and still most widely used representations of information delay (learning) is represented in the well-known exponential adjustment, or Smooth. (Figure 7)

(Figure 7.)



(Figure 8.)

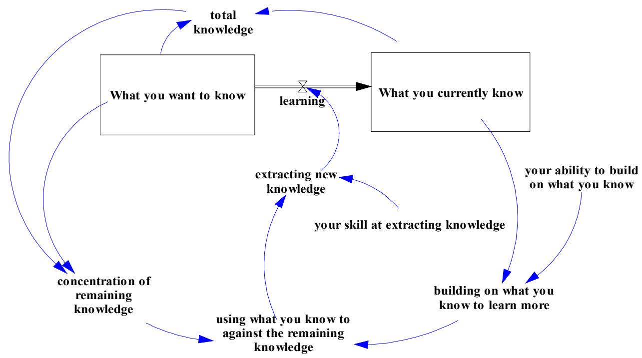


The shape of this 1st order exponential adjustment or Smooth is consistent with the immediate reinforcement of classical Pavlovian Conditioning.[3] Studies in learning theory have show that this type of learning displays the shape of the Smooth. Likewise, “forgetting” displays a curve consistent with a 1st order decay process. For these reasons this type of model is a reasonable representation of one form of learning.

The Learning Curve – Building New Knowledge

Real World learning often does not involve immediate reinforcement; that is, truth is not immediately verified; there can often be long delays in the verification of truth. Additionally, some level of knowledge might be necessary in order to learn faster. This type of learning is more complex and more difficult to represent. The Rescorla and Wagner model of classical conditioning assumes that there is a growing strength in association as learning progresses. The literature reviewed wasn’t all that clear on what dynamics result, however some sources hint this can result in an S-Shaped or sigmoid growth curve[3]. It is reasonable to assume in the case of the analyst that the more one knows the easier it is to associate relationships and spot patterns and thus build new learning faster, or, as stated, the more one knows the faster one can learn. David Ausubel's theories on “meaningful learning” postulate that the most important single factor influencing learning is what the learner already knows. (Ausubel, 1968) Ausubel's theory, perhaps more so than the simpler forms of conditioned learning, clearly implies that an accumulation (stock) of knowledge is causally connected to the learning rate. This would imply dynamics similar to the Diffusion Model (Figure 9) in System Dynamics.

(Figure 9.)



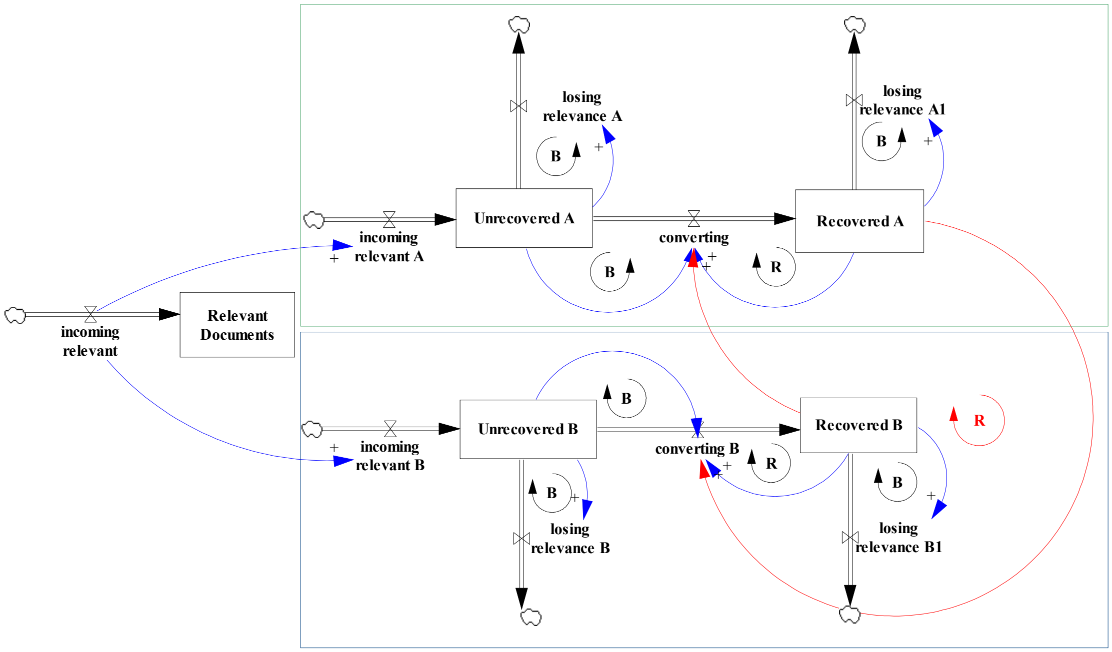
The Limits of Knowledge

There are several problems trying to use the Smooth, or even the basic Diffusion Model to represent the type of learning that I am trying to represent. The basic Smooth contains some of the desired structure, but the goal, or limit, is often a constant. Additionally, the information available and relevant is not infinite, and learning may become more difficult as the information remaining begins to exhaust itself. This would also mitigate against using the Smooth to represent the type of learning I am attempting to represent.

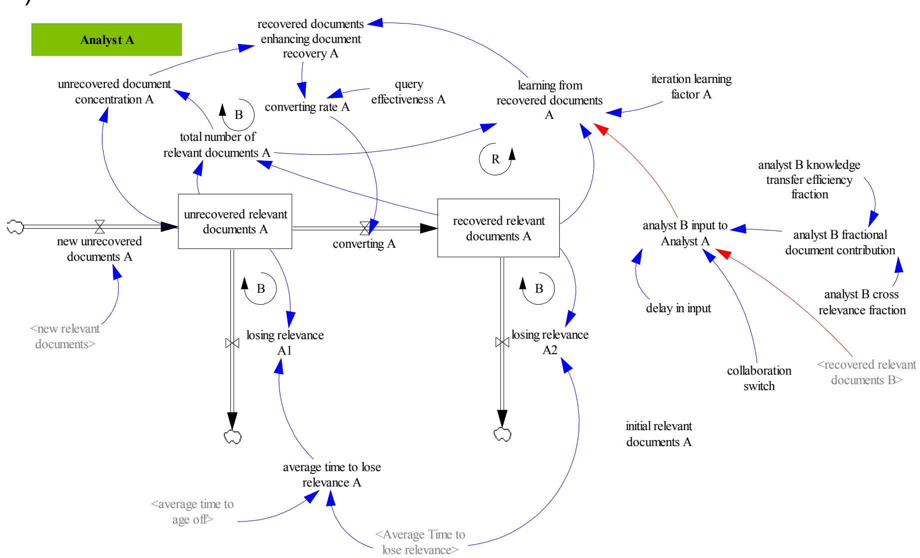
The goal or limit of knowledge changes in the system I am trying to represent; therefore, it is best represented by its own stock. This, in turn, leads us to the Diffusion Model where the goal or limit is represented as a separate stock. The basic Diffusion Model is also insufficient to accurately represent the system. The basic Diffusion Model, although it has two stocks, is technically a first-order system. In the real world system, both what has been learned and what can be learned are constantly changing accumulations; therefore, stocks with inflows and outflows must represent both. Figure 10 shows the basic Stock & Flow structure and feedback loops of the model. The collaborative process is present only if the information arrows and reinforcing feedback loop represented in red are present. A detail graphic of just one analyst's portion is shown in Figure 11 with the collaboration connection again highlighted in red. (See the Appendix for a graphic of the complete model.)

This structure correctly represents the lack of causal connection between converting (learning) and loss of relevance. That is, the rate of learning has on no impact on the sum of the relevance outflows. Put simply: information continues to lose relevance at the same rate without regard to the analyst's rate of learning. Having this effect accurately represented was essential in order for the model to reasonably represent the dynamic properties of the real system.

(Figure 10.)



(Figure 11.)



Collaboration Variables

The addition of the collaboration loop (highlighted in red in figure 10) allows Analyst A to transfer knowledge to increase Analyst B's rate of learning and vice versa. Without this connection the analysts are assumed to be learning on their own, unable to leverage the high confidence knowledge of other analysts. There are several additional variables necessary to represent the collaborative connection.

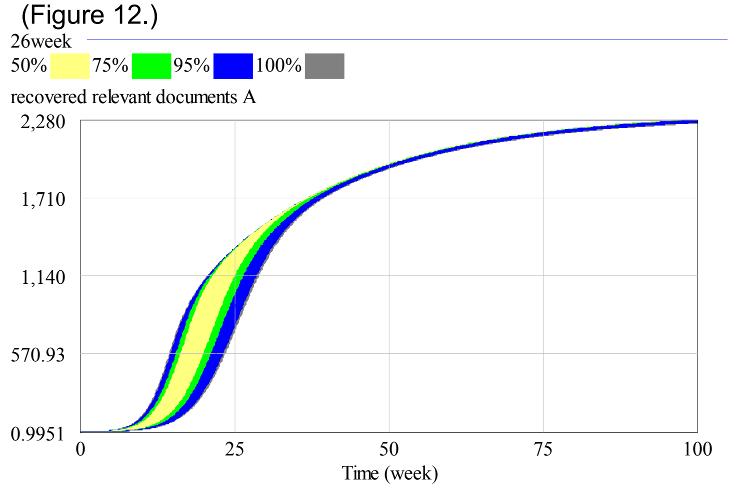
1. Collaboration Switch: On-Off decision switch to share or not to share.
2. Analyst Relevance Fraction: Even Analysts working the same area may assess relevance differently. This variable provides the option to limit cross-relevance to some fraction.
3. Knowledge Transfer Efficiency Fraction: This variable basically weights how effective the analyst is at using his or her knowledge to improve the learning rate of the other analyst. This variable also abstracts several additional factors that affect collaboration knowledge transfer, such as knowledge redundancy.
4. Delay in input: Information does not transfer instantly from one person to another. A one-week smoothing delay is used. This implies that the information is delayed but conserved. Choosing to represent this delay as a separate stock and making it subject to relevance drain would be appropriate in situations where there is a longer transfer delay. I chose to use a Smooth because the additional model complexity would not significantly improve fidelity using the transfer delay I intended to use and retain as a constant.

Testing the effects of Collaboration – Slow & Fast Operational Tempo

After the model has been assembled we can test how different collaboration strategies behave as the time to lose relevance is varied. The Sensitivity Graph shows two analysts collaborating under two different operational tempos. The run comprised 200 simulations varying *knowledge transfer efficiency fraction* from 0% (no collaboration) to 100% (total knowledge to learning rate transfer). The lower right limit to the Figure 12 and 13 sensitivity shows no collaboration, and the upper left-hand limit shows 100% collaboration.

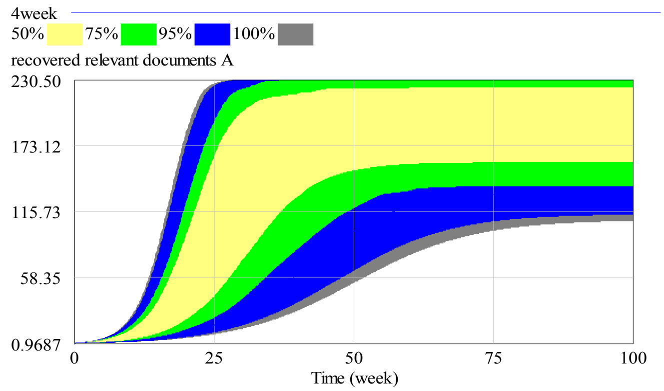
Relatively slow tempo – 26 week "time to lose relevance"

(Figure 12.)



(Figure 13.)

Faster tempo – 4 weeks “time to lose relevance”



Figures 12 and 13 show how effective collaboration is during periods of slow and rapid change in the information environment. Analysts who do not collaborate learn more slowly, and show lower long-term knowledge levels. Collaboration strategies, which exploit the feedback effects of knowledge transfer, show greater success.

However, the results also show that the need to collaborate on problems that are slow moving is more limited. With a long *time to lose relevance* the pace of learning is quickened by collaboration, but because the document set is changing slowly, the capable analyst eventually recovers most of the documents even without collaboration. In contrast, collaboration is almost essential when the *time to lose relevance* is short. That is, the individual analyst even if highly skilled often simply cannot learn fasts enough if working alone. The knowledge transfer from others to increase his or her rate of learning appears essential.

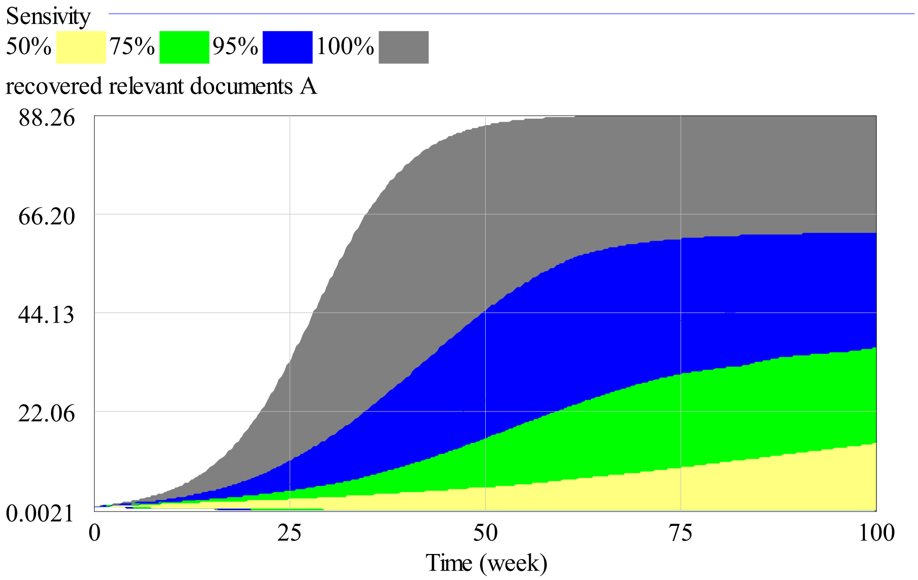
<u>Inside the “OODA Loop”[4] – The Dynamics of Knowing Nothing Relevant</u>

Analysis is normally done to support decisions and decide actions. Decision-making has its own feedback structure, not addressed in the model; however, I would like to comment on how the model dynamics could impact such a decision loop. One description frequently used to describe the decision making feedback loop is “Orientation, Observation, Decision, Action,” or OODA. Under conditions of rapid change there is always the possibility that old knowledge becomes outdated, and that learning cannot keep pace with the rapidity of events. If the operational tempo reaches the point where information changes faster than it can be digested and used there is a risk that no knowledge of any relevance remains. That implies that no effective knowledge-based decisions can be made, or decisions will be made based on outdated information or no information. In essence, learning inflow rate does not keep pace with the knowledge outflow, and the accumulation of relevant knowledge can potentially get drained to near zero. With no apparent valid knowledge to support decisions or actions, decision paralysis often results.

This is essentially the effect the “OODA Loop” was coined to explain. The dynamics implicit in OODA have more application than just the military context from which the concept was generated. That is, it applies equally well to the Boardroom or Analyst Knowledge of a particular problem.

It is possible to simulate the rapid change of the knowledge space (represented by documents recovered) that results in essentially knowing nothing of relevance by simply reducing the *average time to lose relevance.* In essence, *time to lose relevance* and analysts learning rates are in constant competition.

(Figure 14.)



Sensitivity graph varying collaboration efficiency from 0-100% with 2.5 weeks to lose relevance.

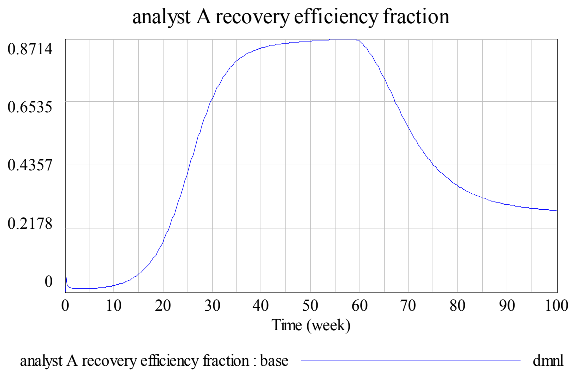
If things are moving really fast there is a significant risk of knowing nothing. In the case above, one-half of the sensitivity output (roughly less than 50% collaboration efficiency) is lost at 0 – essentially the analysts know nothing about the problem if they don't collaborate well![5] Figure 14 shows that individual learning rates are often insufficient. Collaboration, with its information cross-connection and additional reinforcing feedback loop exploits the exponential nature of feedback to enable the collaborative partners to better keep pace with the changing knowledge space.

The knowledge-based necessity for collaboration during a crisis is reflected in the behavior of real systems. It is very common for organizations to generate crash collaborative efforts during a crisis. The real threat of "knowing nothing relevant" also affects the organizational dynamics in analytic efforts. It is very common to gather for a crisis and disperse after the crisis has passed. Collaboration appears to provide a "knowledge advantage" because it more rapidly builds the knowledge stock by exploiting the exponential growth potential of feedback as the analyst pass and build their confidence in the knowledge.

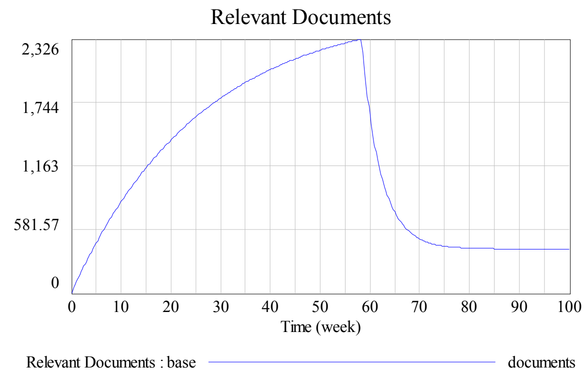
<u>As things change – the dynamics of "Losing the Handle"</u>

A sudden change in tempo can have severe impact on knowledge levels. Valid knowledge (represented in the recovered documents stock) can rapidly erode and the analysts involved may be ill prepared or even unwilling to accept the fact that knowledge they had relied on for years is no longer valid.

A sudden change from 26 to 4 weeks *average time to lose relevance*



(Figure 15.)



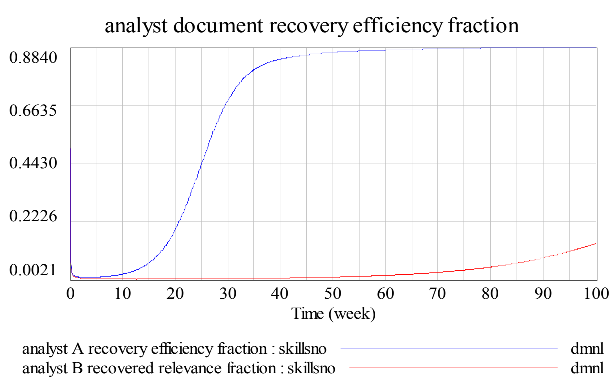
(Figure 16.)

Figure 15 and 16 represent the sudden erosion of valid knowledge when operational tempo changes. Sudden change, such as a Revolution, could cause this type effect. The erosion of valid knowledge can be devastating.

The Effects of Training

Analyst skill affects future learning for that Analyst, but also the future learning of all Analysts involved in the collaborative effort. Most of the Model output in regards to the issue of skill and training corroborate what would be expected from common sense, e.g. a Junior Analyst's learning rate rise significantly in a collaborative setting, but may stagnate in a non-collaborative setting. However, generally, even a Senior Analyst will still benefit from collaboration with very junior personnel.

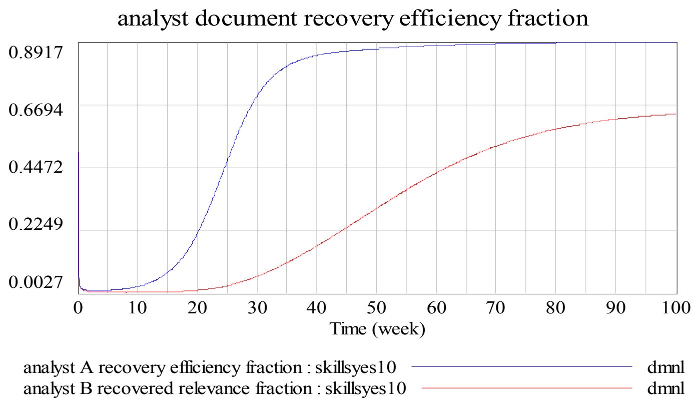
(figure 17.)



A Junior Analyst (Analyst B) left to his or her own devices may never get off the ground. (Figure 17)

The same scenario with Collaboration: 10% knowledge transfer efficiency fraction

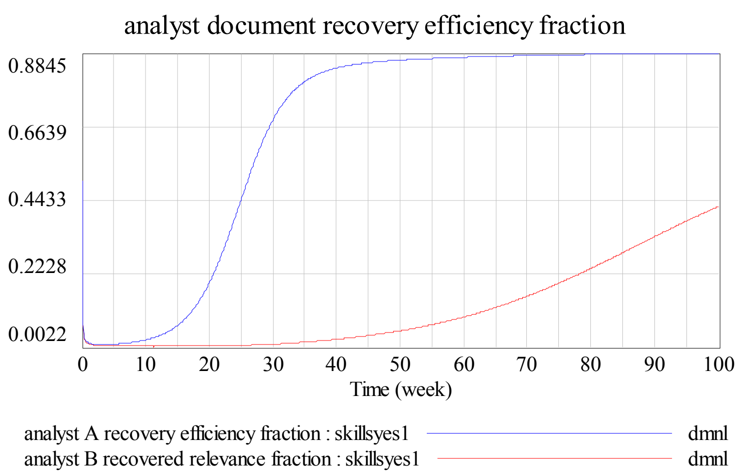
(Figure 18.)



A 10% transfer rate may seem unrealistic; however dramatic gains result from even the most modest transfer.

1% knowledge transfer efficiency fraction

(Figure 19.)

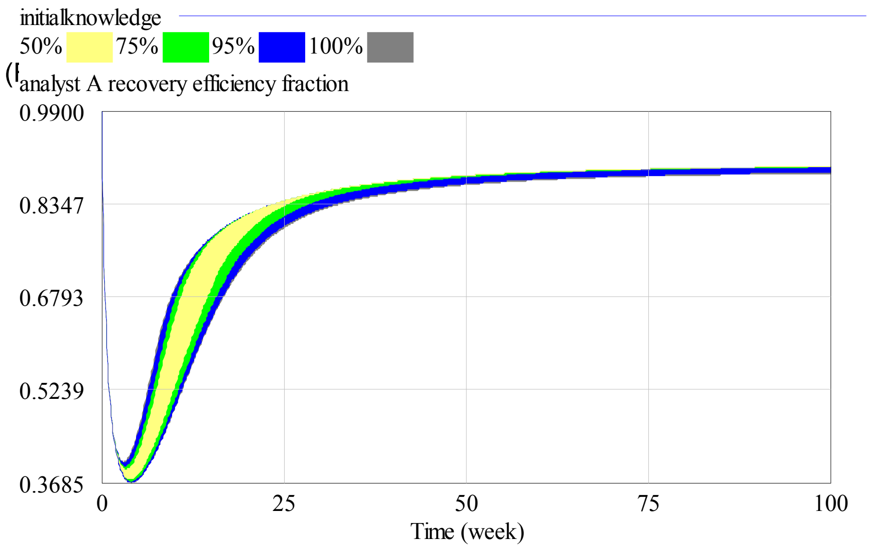


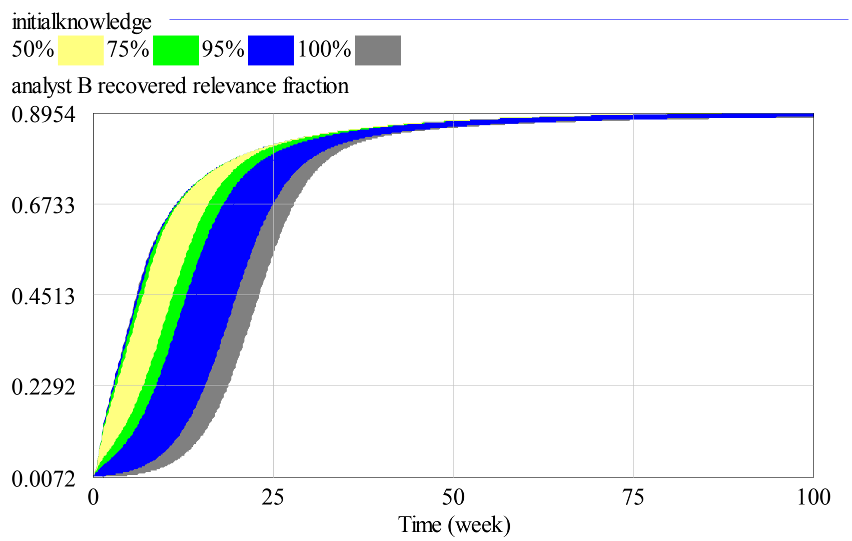
The reasons for these dynamics are – once again – the power of the feedback dynamics inherent in collaboration.

Initial Knowledge

The impact of initial knowledge (defined as number of initial documents recovered) can be significant, but initial knowledge *rapidly degrades in significance as average time to lose relevance declines.*

(Figures 20 & 21)



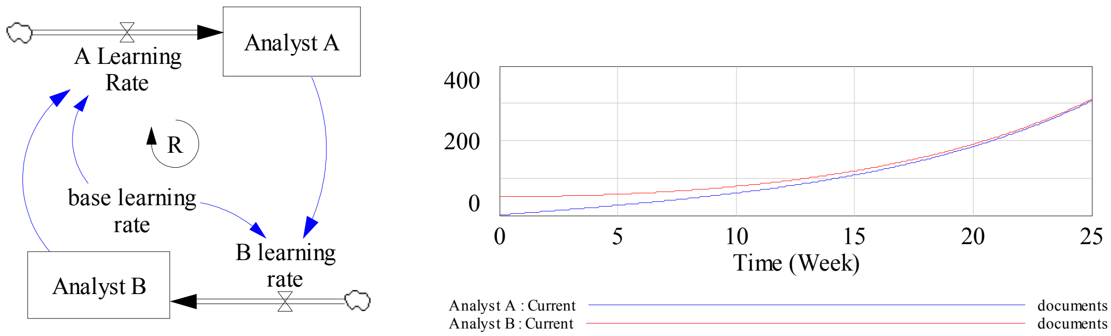


Initial Knowledge and Collaboration: Analyst A has initial knowledge but Analyst B does not. Analyst A has 100 Documents of initial knowledge. Collaboration varies from 0-100%. Analysts have identical skills. (Figure 20,21)

The transfer of initial knowledge decreases learning times even when analysts have identical skill. Even the learning rate for the Analyst with significant initial knowledge is enhanced.

There is at least one more dynamic behavior resulting from the embedded “simple” feedback structures in collaboration that bears mentioning. The behavioral dynamics of high order positive feedback systems (Ashford, 1995) may impact collaborative learning. Figure 22 shows structures, which can show convergent behavior under certain conditions. That is, the values of the stocks can converge over time despite the possible large differences in initial value of those stocks. This would imply that the learning of the collaborative group could converge over time due the dynamics inherent in the basic structures.

(Figure 22.)



Examining the Implications of Feedback Delay on the Decision to Collaborate
Delay is present in any dynamic system, and its impacts often go unappreciated. This, in turn, often leads to faulty decisions and actions because the often subtle effects of delay were not considered. In order to demonstrate the feedback effects that underlie the basis for the effectiveness of building group knowledge through collaboration. I will start by examining the simplest models possible.

Most people understand intuitively that feedback is one cause of the effectiveness of collaboration and the building of knowledge within a group. However, as with many mental models of dynamics, the true implications of those dynamics are often missing, even in the rare cases when dynamics are correctly understood. Numerous publications and studies have observed that most people are unable to correctly evaluate the effects of feedback on system behavior, (Sterman, 2000) and linearly extrapolate exponential growth (Dorner, 1996). While many people “know” feedback aids collaborative learning, few are able to correctly assess the implications, and thus are unable to use that knowledge to generate more effective collaboration strategies.

In our case, adding an information flows between the analysts creates a new positive feedback loop with the potential to generate the exponential growth of knowledge. However the benefits of the collaboration strategy may be delayed because of the nature of exponential growth. In other words the benefits of the collaborative exchange may take time to “take off.” Other strategies may appear to be more productive in the short term especially if the perceived cost of maintaining the collaborative relationship is high.

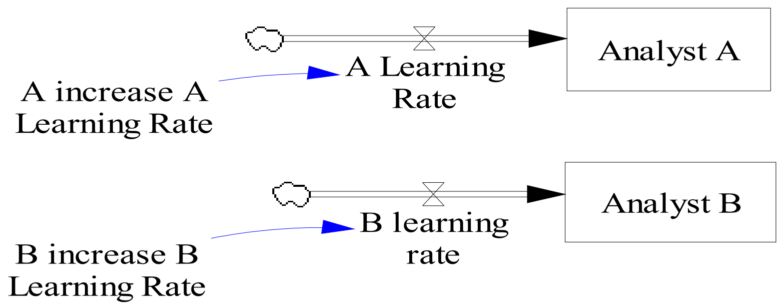
In order to demonstrate these effects in the simplest terms, we pose the following question:

Which is the more useful knowledge building strategy for two analysts?

We increase the individual rate of learning 10%; that is, from the base rate of 10 documents/week to 11 documents/week.
We sacrifice the increase in our own rate of learning, and instead use 1% our knowledge to increase the other analyst’s rate of learning.

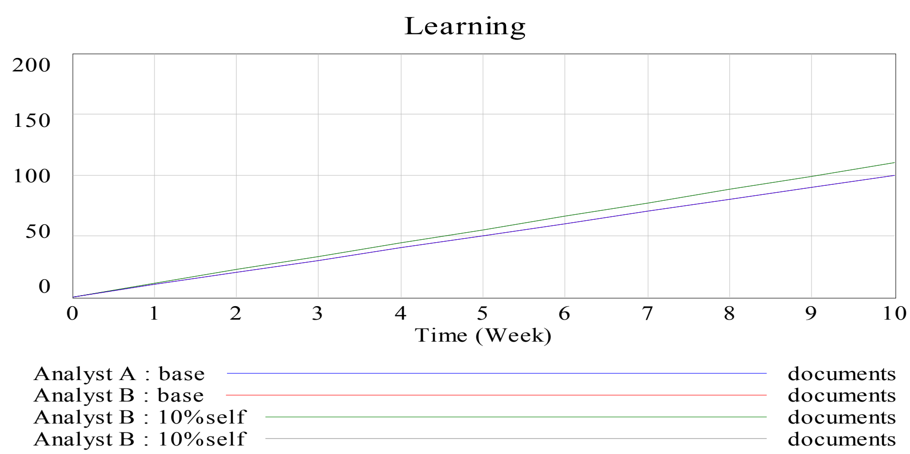
Both Analysts opt to increase their own rate of learning by 10%. (Figure 23.)

(Figure 23.)

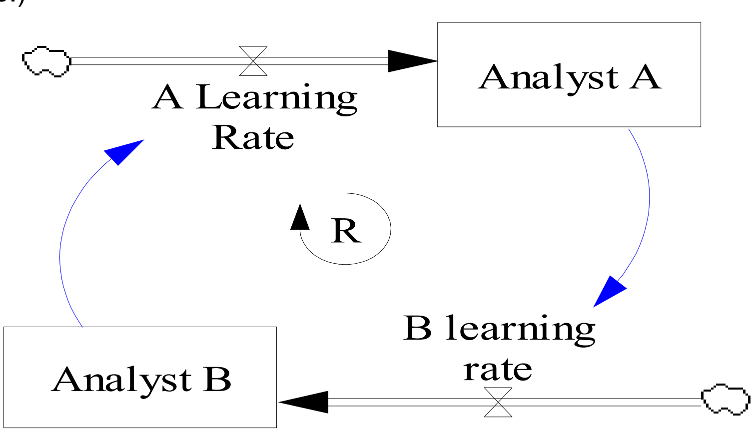


This option leads to the obvious linear increase for both Analysts in Figure 24.

(Figure 24.)



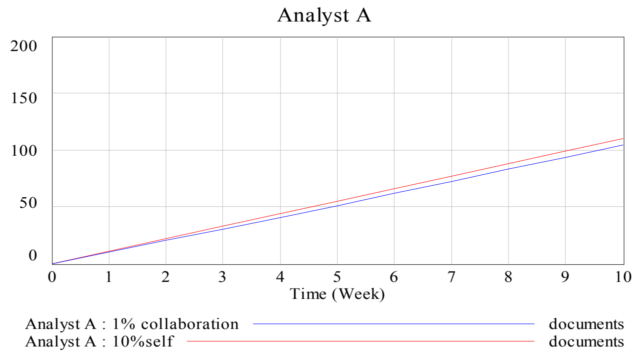
(Figure 25.)



Using a Collaboration strategy instead of a linear individual strategy, both Analysts instead choose to use 1% of their knowledge to increase the other Analyst's rate of learning. (Figure 25.)

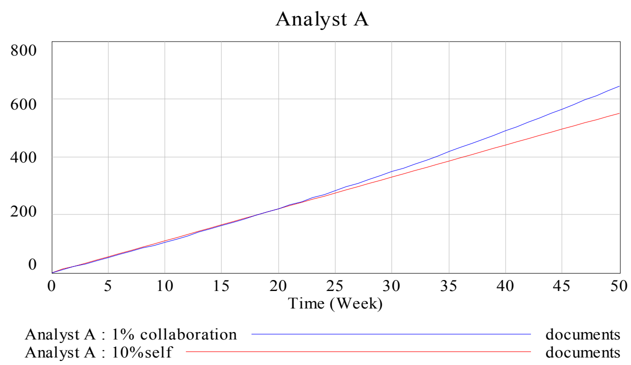
The collaboration strategy *appears* to fail when viewed over the same time period. (Figure 26.)

(Figure 26.)



Viewed over a longer time period a different picture emerges. (Figure 27.)

(Figure 27.)



The choice to transfer knowledge to boost the other's learning shows a dramatic *long-term* impact – due the effects of reinforcing feedback.

<u>Collaboration or Competition – The knowledge trade-off</u>
These simple models only purpose is to emphasis the point that the power of feedback in group learning should not be underestimated. Even with the apparent 10-1 disparity in effort the long-term results favor collaboration.

The fact that the break-even point for learning may be long delayed when the collaboration option is chosen might have some significance. In practical terms, collaboration efforts where the positive effects of the collaborative feedback take a significant amount of time to take hold are less likely to succeed. Impatience or lack of understanding of the power of the feedback process may cause analysts to abandon the collaboration in favor of other strategies.

Numerous modeling efforts (Axelrod, 1997) have documented the choice between cooperation and competition. Individual assessment whether to compete or cooperate could be based on some utility evaluation. Normally, short-term utility benefits competition, while long-term advantage is usually gained more through cooperation. One of the conclusions of this paper is that there are valid knowledge-based reasons to collaborate and not to collaborate under certain conditions. As shown above the decision to build one's own learning often has short-term benefit, while the collaboration option shows may show superior long-term results. This means that if the Analyst were to make some utility-based evaluation, the decision would ride on which time-window was chosen for that evaluation.

Naturally, these small models are almost absurd simplifications. As the saying goes, all models are wrong – but some are useful. *Even in the more complex model, which is the main topic of this paper, the underlying dynamics of feedback are really no different from these simple models.* Often, in the case of fast operational tempo where one person simply "can't learn fast enough" to keep abreast of the problem, the group can exploit these feedback effects so that in many cases their learning can keep pace.

<u>Conclusions</u>

The exact workings of collaboration become clearer when put in an operational System Dynamics context. Collaboration is frequently cited in management and even knowledge management literature as some near magical phenomenon, but it is left to the reader to attempt to discover how collaboration actually *works.* The cobwebs are largely swept away when it becomes clear that collaboration is a feedback process where knowledge *stocks* increase learning *rates.*

Collaboration is a feedback process, where the exchange of high confidence knowledge builds knowledge "stocks" within the individuals of the group and that valid knowledge then increases the learning rates of all. That allows individuals within the collaborative group to keep ahead of the remorseless workings of information relevance decay. It is important to note that individuals must learn separately – the concept of a group "knowledge stock" is a fiction. I found it impossible to establish any meaningful operational definition for such a concept. Some writings acknowledge this (Fenwick, 2002). Group or "Corporate" knowledge is not a database or even some nebulous summation of the knowledge of the individuals possessing magical powers – it is rather the information connections among the group to more effectively use individual knowledge to build learning rates.

These statement appear simple and obvious now; however, to paraphrase Stephan Wolfram – it's amazing how easy it is to miss what will later appear obvious. That which appears obvious at the end of learning often was not obvious at the beginning. To make matters worse – the implications of the obvious, which should be equally obvious – are also often far from obvious!

The decision to collaborate for knowledge-based reasons are can be defined by the operational tempo, or a sudden convergence of goals. The power of collaboration is derived from the effects of feedback, as knowledge is used to speed learning rates. Under ideal conditions collaboration can exploit multiple reinforcing feedback mechanism to generate super-exponential growth.

While there appear to be valid knowledge-based reasons to collaborate under many conditions, there are also valid reasons *not* to collaborate under other conditions. Collaborative efforts always involve some trade-off between the time required to service the collaborative effort and the gains of that effort. Under certain conditions, where the *average time to lose relevance* is extremely long, Analysts may decide that collaboration is not in their best interests. Whether the evaluation of the knowledge-gain utility plays into personal choice to either collaborate or compete is an open question.

There are of course numerous other reasons to make the choices involved in collaboration, such as if collaboration is likely to enhance or degrade the chances for advancement or reward. Such a utility evaluation is common in many organizations. Unfortunately, the perceptions that lead to such behavior can persist over time even when operational conditions have shifted. This can have significant impact and could lead to potential policy resistance to collaborative efforts, when they are most needed.

The apparent requirement to either “collaborate or be at risk of knowing nothing relevant” is strongly mirrored in how people and organizations behave under times of rapid change and crisis.

I did not start out to “prove” collaboration works, but to try to understand it by attempting to define it in the operational sense using a System Dynamics model – largely to reduce my own level of ignorance regarding the workings of collaboration.

<u>Implications for Collaboration Strategy Generation</u>
It is risky generalizing beyond the admittedly narrow boundary and design purpose of the model. However, in the spirit that all conclusions are actually just new and hopefully better hypotheses, I will hazard a couple of observations regarding generating Collaboration strategies based this work. Most strategies for the building knowledge mostly revolve around manipulating data inflow in the “analysis system.” These strategies have many known side-effects, which can result in analysts being drowned in data but starved for insight, or to paraphrase the Philosopher Heraclitus – access to multiple terabytes of information does not imply insight. Data is not knowledge, and knowledge, once built, does not retain validity for all time. Knowledge isn't a constant.

The collaboration literature tends to emphasize the capability of the group to rapidly add new information and the power of the positive feedback structure. This is an incomplete picture of collaboration, and indeed why many collaborative efforts fail. Many failed collaboration efforts result from an under appreciation of the more subtle dynamics of positive feedback, which, for example, can lead to group think, or the group being paralyzed by too much data. What may be lacking in failed collaboration is the essential balancing feedback structures required for effective critical thinking and decision making.

Effective collaborative efforts do concentrate on the exchange of “valid” information, or to put it more precisely, information in which the analysts possess greater confidence. The power of the collaborative efforts that succeed is not necessarily due to adding information, but that information or knowledge assumed to be needed is *invalidated and deleted* from the group's consideration. Balancing feedback structures essential for effective collaboration generate those processes. Unfortunately, those structures are often weak or missing in failed collaborative efforts.

“It is not the (collaborative) network or its robustness that determines the quality of the information sharing. What matters is the nature of our thought processes...” (Reid, 2003). Additionally, while collaboration thrives in rapidly iterating hypotheses based on a variety in points of view, simply bringing in the most undocumented and unpublished mistakes is often enough to avoid critical mistakes!

Other strategies involve trying to build learning rates, often through key word or data mining techniques. In my estimation these well-intentioned efforts often have only incremental or linear impact, often because feedback processes for improvement are weak, very slow, or missing altogether. The key to collaborative learning appears to be the establishment of the information connection for the exchange of human knowledge of high confidence because it has survived critical evaluation, or the group evaluates the information critically as a group. Additionally the group must correctly choose what *not to learn. That*, in turn, can further human learning exponentially. Concentrating on increase the percentage of knowledge exchanged and invalid knowledge removed from consideration has far greater impacts than increasing raw data inflows.

What appears most important of all is that invalid knowledge be destroyed through iterative hypotheses testing in order to retain only the information that is of high confidence and relevant. Successful collaborative groups have a habit for zeroing in on just the knowledge necessary to solve the problem – ruthlessly eliminating unneeded data and weak hypotheses.

<u>Follow-on Work</u>
There are additional avenues for investigation resulting from this effort. More work is possible investigating the impact of simple generic structures on the dynamics of collaboration.

One often under appreciated requirement for real learning is that discarding invalid old knowledge is probably more important than gaining valid new knowledge. As mentioned earlier, the modeling of out-of-date knowledge prejudicing current learning is possible by adding a stock of assumed knowledge, which accumulates at a different rate because the assume *average time to lose relevance* is too long. This causes out of date knowledge to accumulate at a faster rate then valid knowledge. How that affects the current rate of learning is the matter that could be investigated.

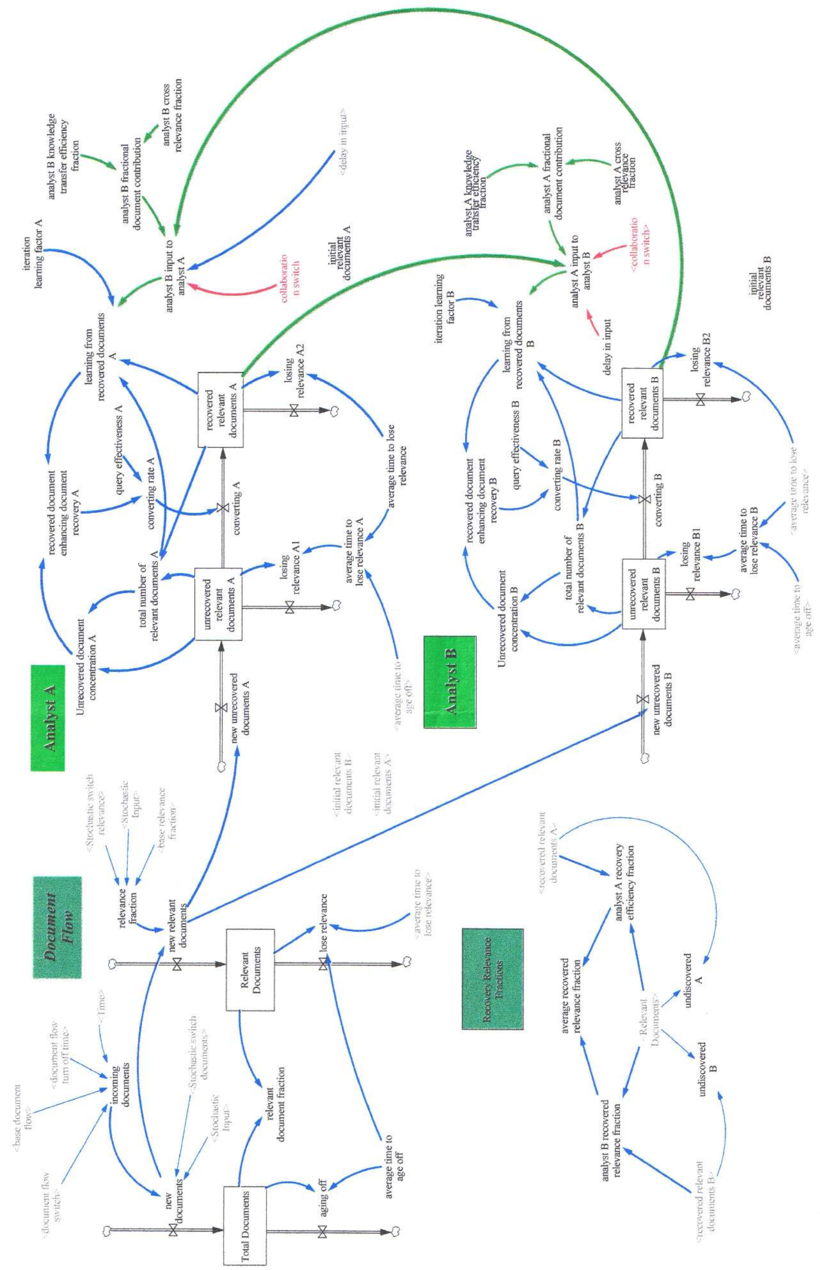
Additionally, small group dynamics is another possible point for follow-on work. Initially, I had strongly desired to avoid pursuing modeling small group learning dynamics, because I didn't think the added complexity was necessary to answer the questions I built the model to investigate. That is why I limited myself to the simple two-analyst model that I eventually decided on.

The initial investigation into small group dynamics using models similar to the one used indicates that modeling small group dynamics is potentially possible. The math of estimating the sum of group knowledge and how that is distributed amongst the group gets messy pretty fast.[6] My current instinct is that a less precise approximation could be used to estimate the sum of group knowledge, and that that would suffice for the level of abstraction already acknowledged in these models.

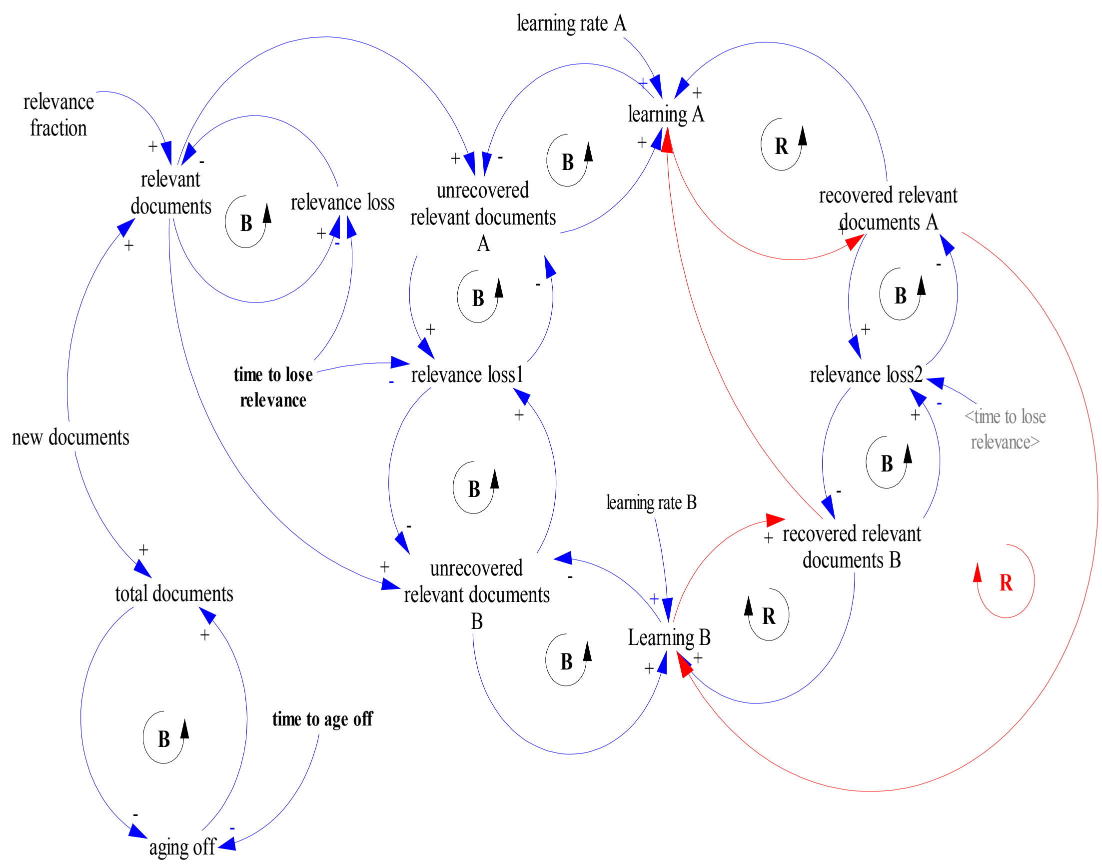
It is also possible to add the decision to collaborate or not collaborate based on perceived need to do so. This could be accomplished by having the collaboration fraction influenced by a variable tracking the analysts perceived knowledge gap. Conklin (2003) suggests using *net interactive value*, drawing on the concept of *net present value (NPV)*, as a basis for evaluating the present utility of the collaborative effort. He does not detail a specific formulation.

Appendix

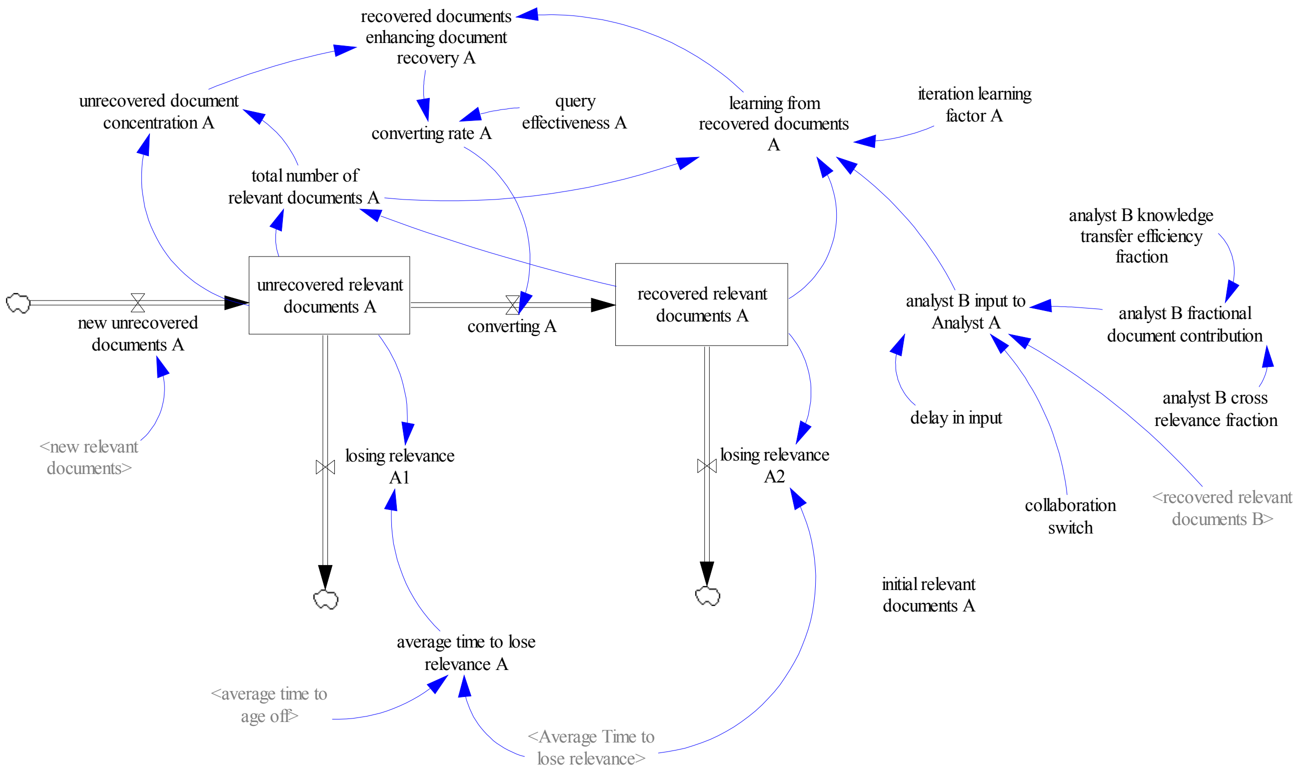
The complete model



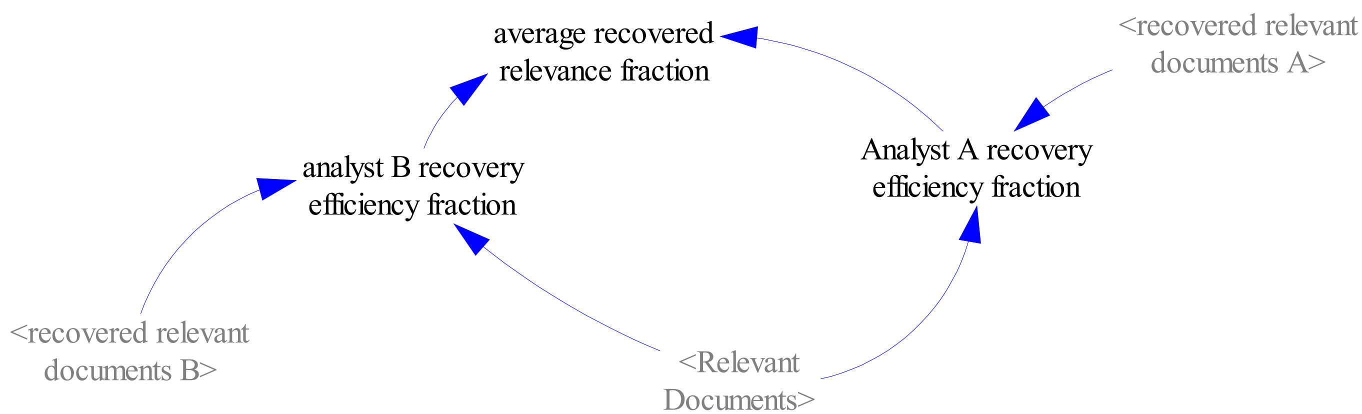
Causal Loop Diagram of Major Feedback Loops



Analyst Learning & Sharing (All) (1 of 2)



Relevance Fraction Calculations



Notes:

1. Relevance - "Bearing upon, or properly applying to, the case in hand; pertinent; applicable." To a great extent relevance is determined by what we already know and usually determines the analytic boundary. The concept of information relevance having some average time constant is an acknowledged simplification in the model; different types of information may have a wide range of how long an analyst might consider the information relevant. There are numerous additional factors that play into how long a particular piece of information is assessed to be, or actually is, relevant to the problem at hand. For example, the most recently viewed information is usually deemed more relevant than older information, even if this is not in fact the case. Examining these issues is not necessary to the purpose of the model, and so is outside the model boundary. Few sources have mentioned this concept in relation to collaborative efforts. Conklin (2003) uses the term "information entropy" in relation to collaboration, but he does not detail how such a quantity could be measured. It should be noted that the concept as defined in this paper is not analogous to the more formal definitions of information entropy proposed in information theory.
2. In the real world the amount of knowledge relevant to a problem probably has real limits – and most certainly has practical limits. Unfortunately, the concept of *time to lose relevance* strays towards metaphysics and epistemology – a valid criticism. This did not erase the need to define some way to calculate some epistemological quantity akin to the "limits of current relevant knowledge." What we choose to define as relevant is constrained by what we currently know or think we know – our mental models. That in turn determines what data we choose. As an aside, and to risk a very broad generalization, epistemology is mostly concerned with what in System Dynamics would be considered the inflow (learning) and what can be contained in the "stock" of knowledge. Much less consideration is given to the outflow; that is, what we forget, considered irrelevant, evaluate as noise, or simply weed out of our thinking processes due to the limits of bounded rationality. Again, to risk a very broad generalization, that portion of the knowledge equation has been addressed more within the realm of cognitive psychology than epistemology.

3. http://books.pdox/Physics/Theoretical%20Neuroscience/reinforce.pdf
   Some discussions of the RW model suggest the true shape is sigmoid
   http://brembs.net/classical/rwmodel.html
4. **Observation**, **Orientation**, **Decision**, **Action** - A phrase representing the military decision feedback loop coined by Col. John Boyd, USAF (Ret.)
5. The effect of invalid "old" knowledge biasing and inhibiting current learning is well known. Representing this is possible by tracking a stock of knowledge that is assumed to be true, but is actually outdated.
6. Messy is a technical term for complex and computationally intensive. The problem of redundant document selection, which was mentioned earlier, is largely to blame. I would like to thank my more math-enabled colleagues for helping in working on the probability aspects of the math after I had described the problem. The current preliminary or initial hypothesis is that group learning stock can be calculated and follows a hypergeometic distribution. ((j matches/redundancy, N docs, n A learning, m B learning} prob=(mCj)((N-m)C(n-j))/NCn. This implies, for example, that 3 analyst learning at a rate of 5 will outperform 5 analysts learning at a rate of 3 – this corroborates the anecdotal efficiency of small skilled collaborative groups. This line of investigation is in a preliminary stage.

References

Sterman, John. 2000. *Business Dynamics – System Thinking and Modeling for a Complex World*

Dorner, Dietrich. 1996. *The Logic of Failure*

Dixon, Nancy. 2000. *Common Knowledge – How Companies Thrive by Sharing what they Know*

Fenwick, Tara. *Limits of the Learning Organization: A Critical Look*, http://www.ualberta.ca/~tfenwick/ext/pubs/print/lngorgeric.htm

Kerr, Norbert L. Tindale, R. Scott, *Group Performance & Decision Making,* Annual Rev of Psychology, 2004 55: 623-55

Axelrod, Robert. 1997. *The Complexity of Cooperation*

Davis, Morton D. 1983. *Game Theory – A Nontechnical Introduction*

Senge, Peter. 1990. *The Fifth Discipline*

Senge, Peter et.al. 1999. *The Dance of Change*

Linden, Russell M. 2002. *Working Across Boundaries – Making Collaboration work in Government and Nonprofit Organizations*

Neustadt, Richard. May Earnest. 1986. *Thinking in Time – The Use of History for Decision Makers*

Popper, Karl. 1957. *The Poverty of Historicism*

Popper, Karl. 1934. *The Logic of Scientific Discovery*

Ausubel, David. 1968. *Educational Psychology, A Cognitive View*

E. Jeffrey Conklin. 2003. *Wicked Problems and Social Complexity*

E. Jeffrey Conklin, et. al. 2003. *Towards an Ecological Theory of Sustainable Knowledge Networks*

Giffen, R. Reid, E. 2002. *A Woven Web of Guesses: Canto Two – Network Centric Warfare and the Myth of Inductivism*

Ashford, Araron C. 1995,2001. *Unexpected Behaviors in Higher-Order Positive Feedback Loops (D-4455-1)*