Normann's Crane 2.0 Operationalising Service Systems co-design

Richard Taylor and Chris Tofts
Trusted Systems Laboratory
HP Laboratories Bristol
HPL-2007-160
10th October, 2007*

service system;
co-design; model; SLA

Service development and delivery is by its very nature an activity that requires cooperation between suppliers and customers. This cooperation does not imply or require parity between the parties. In any such relationship parties divide responsibilities based on a perception of self-competency and subsequent economic value. This is not always rational. One result of many such co-production agreements is to leech competency to one partner to the detriment of *all other* parties. Black box relationships are an unintended consequence of such interactions, relying on an ability of all of the partners to express their requirements in terms of rigid service level agreements whose origins are opaque to all but the originating organisation (and even then may not be well understood). This unhappy position is made worse by a tendency to involve many service providers at different points in the provision chain, none of whom have an adequate or rational understanding of the requirements of their customers or suppliers.

This paper argues that true co-production can never be successfully met without effective co-design - achieved through the deliberate exposure of internal structure and imperatives by all parties. Apparent productivity gains will be short lived without this exposure as organisations limit their choices of partners in markets as well as restricting their own ability to address new opportunities.

---

* Internal Accession Date Only

Normann's Crane 2.0
Operationalising Service Systems co-design

Richard Taylor and Chris Tofts
{richard.taylor, chris.tofts}@hp.com
Hewlett Packard Laboratories
Systems and Services Research Centre

Abstract

Service development and delivery is by its very nature an activity that requires cooperation between suppliers and customers. This cooperation does not imply or require parity between the parties. In any such relationship parties divide responsibilities based on a perception of self-competency and subsequent economic value. This is not always rational. One result of many such co-production agreements is to leech competency to one partner to the detriment of *all other* parties. Black box relationships are an unintended consequence of such interactions, relying on an ability of all of the partners to express their requirements in terms of rigid service level agreements whose origins are opaque to all but the originating organisation (and even then may not be well understood). This unhappy position is made worse by a tendency to involve many service providers at different points in the provision chain, none of whom have an adequate or rational understanding of the requirements of their customers or suppliers.

This paper argues that true co-production can never be successfully met without effective co-design – achieved through the deliberate exposure of internal structure and imperatives by all parties. Apparent productivity gains will be short lived without this exposure as organisations limit their choices of partners in markets as well as restricting their own ability to address new opportunities.

Introduction

In his seminal work 'Reframing Business' Richard Normann (Normann 2001) outlines the properties of a *crane* that enables the observation of the requirements for achieving the reframing of a business. In summary he lists the following properties of the 'crane at work'

- it must start as Gregarian (Dennet 1995);
- emergence is fundamental;
- a necessary combination of scientific method with open ended imagination;
- sourcing from memories in three dimensions;
- allow dialectic between distance and focusing;

- the reduction of dramatic reductionism of the human brain;
- it must be social participative, interactive and include action learning to socialise, encode and create tacit knowledge.

Over the last five years, the authors have developed and deployed methodologies and processes both within Hewlett Packard, as well as with its customers and partners that have addressed many of these issues. From our experiences a further essential property of any ‘crane’ operation in this environment is that it can shift the mental objects rapidly enough to reflect the necessary systems requirements (observation and control). Whilst attempting not to be prescriptive, Normann leaves open the question of what is embodied within each of the activities or requirements. As a consequence, the primary, only, mode of discourse that would be employed in solving all of these problems will be natural language. In todays business context this will be English. Whilst the primary deployment of English as the only mode of discourse is very helpful for a North Atlantic cabal, it necessarily limits the construction of the *view of many heads* as espoused by Normann himself.

The use of natural language as a primary mode of discourse for describing complex interacting systems has a long and dishonorable history in the production of computing systems that do not function as desired. In fact using natural languages approaches in these complex interacting environments almost never leads to a correctly, or even satisfactorily functioning system.

Whilst for those with a particular training replacing the *politics of the spreadsheet* (Schrage 1999) with the politics of the elegantly expressed argument may be of great value to them. This does little for either identifying or resolving the inevitable trade offs between competing stakeholders. In fact it is extremely hard to make a rational decision on a trade off that is neither identified or quantified. One of the primary factors we found limiting the success of complex services projects is the ability to understand the dynamics of the service in operation. The move from the ‘*photo to the movie*’ is a difficult one for any human to make. As observed by Normann, success is not achieved just by having a ‘fit’ today. How ‘consonance’ or alignment is achieved and maintained with the environment, within internal stakeholders and between internal and external stakeholders is made more problematic when the reasons as to why it was ever there or likely to maintained were never made explicit. Much as in the Toyota production system, there is a question as to how much and what needs to be shared between internal and external shareholders in order to maintain ‘consonance’ going forward. The abscence of any real understanding of the joint system dynamics and its rational with respect to the external environment makes the design and maintenance of the appropriate information exchanges at best a guess, at worst utterly incapable of addressing the long term alignment.

Normann correctly identifies that double loop learning is the essence of maintaining success in the current highly dynamic business environments. Unfortunately, like most of the above there is little specified on what the primary objects of this double loop learning should be. Drawing inspiration from the scientific community (even if this means shifting from the from Gregarian to Popperian world view), the primary objects of the second loop of the learning should be the models which lead to the internal loop optimisation decisions. The belief that these models have to be expensive, or even impossible, to produce and maintain is one of the major handicaps of the *post Taylorist* (Miller 1993) view of business (Taylor and Tofts 2003). With this view the Gregarian position is one of triple loop learning, how should we write our models?

The co-design challenge

It should be obvious that no supplier should provide an artifact or service to a customer without understanding the specification of the object that they are meant to be providing. In the manufactured world this will typically take the form of a *functional* interface (what information and/or control passes across the interface) and a *non functional* interface (how fast it has to happen, with what reliability and at what cost). The success of this approach relies on a tacit understanding between both parties of the objective of this relationship. In the case of a television for example, the role of the video encoder can be adequately understood in terms of performance, power consumption and quality (three typical properties). Meta considerations (an expected reduction in cost, power consumption, supply flexibility, licensing and international standards conformance) are retained as tacit agreements outside of the technical specifications of the original agreement.

Services remain more complex. What are often tacit or commercial adjuncts to the supply of physical (hardware) or logical (software) contracts need to be embedded within an explicit agreement between the engaged parties. To take a trivial example from managed information systems – an agreement to provide an electronic mail service with a specific storage capacity, delivery schedule and availability is straightforward. Specifying the impact of requests to scale that service to meet changing requirements (up or down, performant or functional) is non trivial and requires a level of coupling between business requirements, business futures and service provision that is not necessarily part of a natural discussion on the boundary between customer and provider.

A fundamental problem with many approaches to this problem is the acceptance that the point of interface – typically the service level agreement – is not only an appropriate place to measure and enforce commercial agreements, but should be the starting point for understanding the reasons why that interface has been established in the first place. Typically some higher level discussion of the 'why' has been used to establish the fundamentals of the service level agreement, but that is rapidly lost as a service programme evolves (Taylor and Tofts 2005).

This is not a new problem. Organisations typically operate through mixed levels of abstraction and necessarily detail. Information is shared between layers of management (specification) based upon a perceived need to know and understand which naturally flows down the abstraction chain (from business objective through implementation). As requirements flow down the chain, the 'why' is lost, making it difficult for individuals or organisations to feed appropriate information back up as to the impact of low level decisions on overall systems effectiveness.

A failure to provide a means of reasoning about the connected 'why's', even without compounding that with some for of tracking the 'evolving why' lies at the root of many large service failures. Take one example - the ill fated Working Family Tax Credits (WFTC) service, introduced in the United Kingdom in 1999. This service was established to replace the existing Family Credits system with the intention of encouraging people back to work through a tapered income support, operated by and paid through the existing taxation system. In 2001, more than one million families within the UK were reported as benefiting from this support.

The implementation, management and evolution of this service is widely perceived to have been shambolic. Overpayments based on outdated income information (which are often recovered at great expense), underpayments (due to an inability to assess and/or process claims in a timely manner) and frequent delays have discredited what at first sight appeared to be a sensible and reasoned solution to a long standing problem.

The primary issue with the WFTC is one of interacting systems and services. Essentially the system (the service and its interacting stakeholders) consists of five primary interacting sub systems (figure 1);

1. the families who are benefiting from the tax credit – a group whose members are, by definition, often near or below the poverty line, often with young families, and critically dependent on the supplement to an income that often varies seasonally, possibly monthly or even weekly;
2. the support service, consisting of both call centre staff and highly trained field workers who both assess need and advise families on how and when to claim credits. Working dynamics for these groups are dominated in the short term by the time taken to develop, test and deliver scripts to tele-workers (orders of weeks)
3. the computing support service, providing both front and back office facilities and support to workers;
4. the taxation system (with a 'natural' frequency of a year);
5. an the legislature who react to events with frequencies that are dependent upon election dates, news media and constituency feedback.

The fundamental misunderstandings that arise from this approach result from a black box approach to interfaces between the sub systems that inhibits the flow of 'whys' both forwards and backwards as some form of feedback control (Taylor and Tofts 2003a).

This behaviour is even observed within organisations that share a common purpose. Co-production, involving groups who have a (sometimes) tenuous relationship based on hard contractual principles tends to amplify the worst aspects of information sharing (or a lack of it). Informal social relationships within organisations (which often mitigate against some of the more obvious pitfalls) fracture, and the natural fall-back of formal, measurable agreement, quite possibly with no rational base, is used to paper over the cracks.

Co-design attempts to break down the significant barriers between organisations, exposing assumptions that organisations are making in their establishment of requirements and service, providing a common currency for discussion and negotiation, and a base for change management.

Organisations are naturally wary of exposing their internal requirements – now and in the future. By definition, any organisation that has outsourced 'non-core' functionality has defined and determined to protect their core. Revealing these dynamics exposes the raw heart of an organisation. The intellectual property of many 'non' technology focussed organisations relies on this heart. Naturally this makes them nervous – exposing IP accelerates loss of value within their organisations and increases the need to innovate and deliver that innovation.

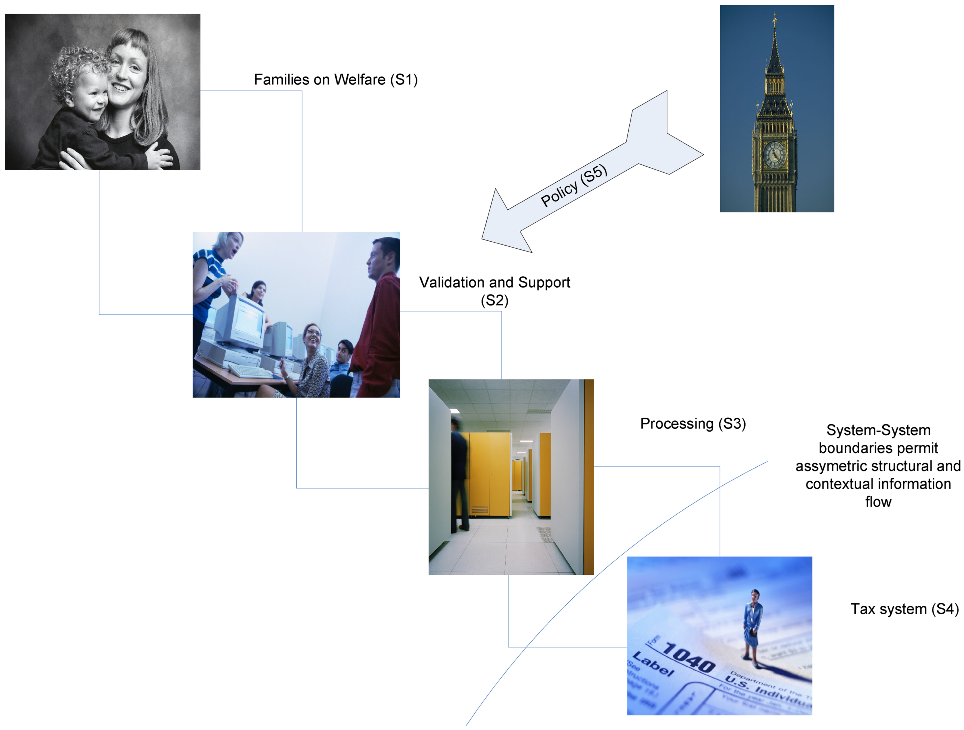

Figure 1: Co-operating systems within the WFTC service system

Co-design through human processes

There is a tendency within humans confronted by a complicated system to concentrate on what is happening - its functions. The immediate goal of a human faced with such a problem is to reduce the overwhelming volumes of data to something they can comprehend, the immediacy (and indeed effectiveness) of human reductionism. One of the primary information losses is that of change. It is difficult enough to capture the functioning of a service system, so how on earth does one capture the dynamics of change? If instead, the process starts by focusing on the areas of trade offs between the stakeholders of the service system, then both the required functions and the implied dynamics are emergent properties of the trade offs. Since any service setting is inevitably one of co-production, the parties are usually aware that there are trade offs between their multiple positions. The consequent human processes require that these trade offs be revealed and expressed within a common language.

The first part of this activity requires that the various stakeholders identify their own issues with the service system and then take part in a meeting where the problems are expressed from each of the stakeholders viewpoints. The second part is the creation of objects, explicit and external to the various stakeholders which express these trade offs. Whilst one temptation would be to embody this object as a form of joint report, a more productive approach is to capture it in the form of models.

In the authors experience, once a list of service system problems have expressed by various stakeholders, it is a relatively straightforward (although obviously **very** highly skilled) process to capture these trade offs in simple models. The plural of models is important. Ownership and buy in (socialisation) of the results of any learning process are vital. One reason for the lack of faith in models of business dynamics is their complexity and distance from any 'normal' business decision maker. As part of this socialisation the authors have demonstrated that leaving small 'errors' or oversights in the models for the participants to correct has a disproportionate effect on 'buy in'. The models then allow the 'future' to be explored both by varying the parameters within the models, or by appropriate changes to the structure of the models.

When the various participants are not from the same organisation, actually, usually even when they are from the same organisation, this activity amounts to opening up the key parameters of service requirement and delivery. This move from a functional black box, *we will send you a request for work and it will be done*, to an understanding of what that will entail for other participants within the service system has a clarifying effect on both the requestor and the deliverer. Discussions as to why certain trade offs have to be made in certain regions of the whole solution space usualy identifies the constraining technical, social and political issues which lead to particular solutions. With this information to hand it is possible to explore what has to change within the service system design space in order that a particular service can function effectively.

Within HP we call the process of obtaining these trade-off positions, and embodying them as models *Rapid Scenario Planning* or RaSP. This approach arose from the authors examination of multiple internal service projects. One consequence of this work was that the construction of internal 'grey' as opposed to black box thinking lead to considerably better decision making and the fundamentals of this approach have been widely adopted within the relevant parts of the organisation. The RaSP process requires the presence of various key stakeholders, a moderator and a modeller. We proceed by asking the various participants to identify and describe their issues well ahead of the meeting, preferably with a 1 month lead time. The moderator and modeller identify the major trade off issues and check these descriptions with the intended participants. Before the main RaSP meeting, 'straw' models of each of the trade off issues a produced and their implications discussed within the meeting. One of the main reasons the models have to be small is that as issues are raised in respect of the trade offs we expect to change the models in the time frame of that part of the meeting. The desire to keep up with the rate of change of the model requirements in a 'live' meeting is an interesting limitation on the approach to modelling that can be undertaken.

One further consequence of this approach is that it forces simplicity and clarity, consequently we have rarely had any issues of comprehension with the models produced. Once each of the trade offs has been discussed and updates to models (including the production of new ones) made - then a final report including models is distributed. Again, the models are distributed in a form where they can be used by the stakeholders. In fact, even when this is a discussion between HP and a customer, the stakeholders are positively encouraged to use the models to locate better 'sweet' spots in the design space. This is in immediate contrast to the 'standard' expert approach to design, where the expert provides the customer (victim) with the parameters of the design they require, with little or no external justification as to why these are the correct numbers, or indeed why other solutions are worse. Whilst many co-producers may wish to live in ignorance it is unlikely that that choice will lead to robust sustainable value producing service systems.

Having captured the service system via a collection of trade-off models it is usually the case that the intended evolution of the system into the future is not just the responsibility of a single stake holder. As a trivial example, a computer's availability is a function of the length of time it can run between breaking and the amount of time it takes to fix once broken. The time between failures is a function of the system design, the time to fix is a function of the design and the service contract. Intriguingly, one can have (not just imagine) systems which break incredibly rarely due to the superlative design and quality assurance. Unfortunately when they do break, the complexity of the design makes repairing them exceptionally difficult, consequently time consuming. In this instance the availability of the system may actually be reduced by the inclusion of more reliable components. A failure to understand the co-design aspects of availability can lead to some very poor design decisions.

Even in this simple setting it is often the case that the system's designers simply see their role in terms of maximizing the time between failures, and the system's maintainers see their role as attending and commencing work on the failure as soon as possible. By understanding the interaction of these key parameters and their likely evolution into the future it is possible to make service design decisions which are significantly more likely to work.

Conclusions

There is considerable agreement that services are complex systems which contain a large degree of feedback. The primary point of departure between communities who study and solve these problems is as to whether complexity should be treated in the same fashion as chaos?

A frequent stop point in modelling exercises occurs at the point at which the modelling team determine that their system is chaotic and consequently abandon any further analysis. The underlying processes of the tidal system for example are chaotic. Thankfully this is on the mega year time scale so predicting water height (up to global warming) at any of the world's ports for the next two centuries is straightforward (there are cheap programs which do it). A similar behaviour is seen with the label complex - the tendency to assume that once something is labeled as such then any need or requirement for understanding is obviated. Consequently the question is how do we understand service systems, in particular given the temporary nature of fit and how does that understanding propagate into the future? In the absence of an ability to explicitly ground the collective understanding of why the service system is and should be successful, the only likely outcome is long term failure. The one way that is guaranteed not to lead to an understanding of service systems is to start with the assumption that they lie within a class of systems that cannot be comprehended.

Equally, the absence of understanding of the trade offs within the system mean that gains in effectiveness or productivity in one part of the system do not play out as gains to the whole system. True productivity and value gains come from the co-exploitation of innovation, that is provision of benefit to multiple parties within the service systems. In the absence of meaningful understanding of the service system it is unclear how the partners would observe and value innovation. In this setting is hardly surprising that the rate of productivity gain in services is often small, and the durability of such gains is short. Given that effective exploitation is inevitably a co-evolutionary problem, co-control is a necessary requisite for success. The questions as to what constitutes effective reward and measurement mechanisms (a marketplace) are largely unanswerable in the absence of an understanding of how a service system should function today and tomorrow. In any economic system the long term participation of all of the parties results from the understanding that there are greater returns together than apart. If there is no understanding of this situation, merely a hope, then jockeying for position with respect to any customer is the inevitable and ultimately self defeating outcome.

Moving from an SLA determined black box approach to service design to at least a 'grey box' co-design approach has clear benefits. However, as many authors have observed, the social processes of arriving at a successful co-design, even within an organisation, let alone between organisations, are extremely difficult. We have successfully deployed the RaSP approach to complex service understanding both within HP and between HP's customers and partners (including competitors in multi-sourced services). Whilst this has required that all of the organisations be more open, it has led to a considerably greater clarity on the current and future goals of all of the participants, and a better understanding of how service consonance is achieved today and will be achieved into the future.

In almost any setting it is far too easy to confuse correlation with cause. In the modern setting with the ability to measure any *dematerialised* process to death, it is far too easy to permit service design to become 'deploy and measure'. The fundamental problem with this approach is that whilst you can measure something and the value obtained may be other than that desired, it does not follow that the value can be changed to a desired one. Treating service design as *deploy, observe performance gaps, and then close them* ignores the problem of prediction.

All of these systems are artificial in the fullest sense. They are solely human artifacts, designed by humans, and embodied by humans and human designed machines. There is no need for them to operate within an impredictive space. However, the absence of an effective design methodology, which extends beyond the logo and the brand name the customer experiences implies and inevitable disappointment in the performance of almost any complex service system.

References

Dennett 1995 Daniel Dennett, Darwin's Dangerous Idea, Simon & Schuster, New York, 1995.

Miller 1993 Stephen Miller, From system design to democracy, Communications of the ACM, Volume 36, Issue 6, 1993.

Normann 2001 Richard Normann, Reframing Business, Wiley, 2001.

Schrage 1999 Michael Schrage, Serious Play. Harvard Business School Press, 1999.

Taylor and Tofts 2003 Richard Taylor & Chris Tofts, Modelling, myth vs reality, map vs territory, HP Laboratories Technical Report, HPL-2003-246, `http://www.hpl.hp.com/techreports/2003/HPL-2003-246.html`, 2003

Taylor and Tofts 2003a Richard Taylor & Chris Tofts, Business as a Control System the essence of an intelligent enterprise, HP Laboratories Technical Report, HPL-2003-247, `http://research.hp.com/techreports/2003/HPL-2003-247.html`, 2003.

Taylor and Tofts 2005 Richard Taylor & Chris Tofts, Death by a thousand SLAs: a short study of commercial suicide pacts, HP Laboratories Technical Report, HPL-2005-11R1, http://www.hpl.hp.com/techreports/2005/HPL-2005-11R1.html, 2005