

Quality of Service in the Long Tail: Narratives and the Exploitation of Soft Metadata

Auke Breemen

► **To cite this version:**

Auke Breemen. Quality of Service in the Long Tail: Narratives and the Exploitation of Soft Metadata. Kecheng Liu; Stephen R. Gulliver; Weizi Li; Changrui Yu. 15th International Conference on Informatics and Semiotics in Organisations (ICISO), May 2014, Shanghai, China. Springer, IFIP Advances in Information and Communication Technology, AICT-426, pp.22-31, 2014, Service Science and Knowledge Innovation. <10.1007/978-3-642-55355-4_3>. <hal-01350907>

HAL Id: hal-01350907

<https://hal.inria.fr/hal-01350907>

Submitted on 2 Aug 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Quality of service in the long tail: narratives and the exploitation of soft metadata

Auke J.J. van Breemen

ICIS , Radboud University Nijmegen, The Netherlands
a.vanbreemen@science.ru.nl

Abstract. The fuzzier domains of social life come into the scope of information science. Below a method is presented that deals with exception handling or also the long tail. It is based upon the Peirce inspired KiF-model. But this time the goal is not to describe procedures or to model elicitation processes, the goal is to develop meaningful sensors that at the same time structure the analysis of the problem and facilitates the access of a database filled with stories. The domain of application is the Dutch educational system.

Keywords: Narrative method, Hard metadata, Soft metadata, Information architecture, Knowledge in Formation (KiF)-model, Education.

1 Introduction

Henry Ford felt the urge to state that any customer can have his Ford model T painted any color that he wants so long as it is black. He wanted to build a car for the great multitude: “large enough for the family, but also small enough for the individual.” Within a century Anderson felt confident enough to state “Forget squeezing millions from a few megahits from the top of the charts. The future of entertainment is in the millions of niche markets at the swallow end of the bitstream.”¹ Subsequently, within a decade, he extends his view from the long tail of the bit-market to the long tail of the thing-market [1, p.63] and resigns from his post as editor in chief of Wired Magazine in order to exploit the long tail in the realm of things himself.

The concept of the long tail is not only applicable to the market. Also with regard to services in the public sphere we meet the phenomenon that about 20% of the cases generates 80% of the classes. In health care the incidence of the more than 14.000² coded diseases diminishes rapidly if we go from the more to the less frequent. In educational and child care services,³ where diagnostics and classification are perused in a far less serious manner, we must assume a long tail, but only some niches are well

¹ See <http://changethis.com/manifesto/10.LongTail/pdf/10.LongTail.pdf>, C. Anderson, The long tail, 13 December 2004.

² For the count see the International Classification of Diseases, ICD-10, 1990.

³ My background is the Dutch educational system, child care and health are overlapping domains. The government decided to introduce Passend Onderwijs (Fitting or Inclusive Education). If this program is taken seriously, we are in dire need to get a hold on the long tail in education.

defined, the majority is crudely defined, if defined at all. A major contributor to the complexity is the circumstance that like personal traits, lead to different results due to the effect of circumstances and experiences on the developing personal character. This tail is besides long, very messy. Oftentimes the stakeholders disagree about the policy that ought to be effectuated. This is a result of not defined or only vaguely defined niches in combination with differences in goal orientation.

Whereas Ford could, in principle, specify every detail of the model T production process in sufficient detail, in the long tail of services a less directive approach is called for. The most we can aim at is that in the long run the classes get sorted out better and adequate policies are suggested. The solution I present below aims at just that. It consists in a narrative approach that utilizes hard (facts) and soft (perceptions) metadata for retrieval purposes. The soft metadata are part of an information architecture that helps the user to find relevant information in other stories told and, according to Kalbach an important requirement for information systems that cover niches, it makes sense by providing context [2]. Interestingly, the model together with the soft metadata proved very useful in the actual handling of long tail cases, see section 2. This offers the possibility to integrate the construction and utilization of domain specific information systems in regular workflow processes, see the conclusions section 4. For the sake of efficiency, in section 3 after the introduction of the proposed approach, I will indicate some differences with affiliated research.

2 A case study: narrative research in the educational system

A disadvantage of strictly quantitative research in the services sector is that although one gets informed about the numbers of people that occupy the different categories distinguished, and about numerical changes through the years, one doesn't learn anything about the diverse dynamics that did lead to the membership of particular categories, nor is there any information about class hopping of particular individuals throughout their careers. As a result this kind of research is of little use for interventions in individual cases. If we want to improve on quality of service we are better off using dedicated information bases filled with Geertz' thick descriptions [3], Ginzburg's micro stories with clues [4] or Schank's scenario's and scripts [5], dealing with individual cases that are (partly) similar, than with quantitative population survey results. Although of course the combination of both is most fruitful and necessary for policy making in the system under scrutiny.

In 2012 Van Garderen of Top Innosense b.v. and I set off collecting stories with hard metadata about pupils that fell out of the educational system.⁴ After some twenty five stories we organized, in cooperation with Gedragswerk,⁵ a meeting with the respondents and key professionals in the problem area. Our main objective was the

⁴ In the Dutch system this is regarded a criminal offence that instigates the officials to start prosecution regardless of the reasons to drop out, which explains why predominantly parents told their stories.

⁵ Gedragswerk is an initiative of the ministry of education. If asked by parents, the organization delivers 'sparring partners' if the parents have conflicts with school regarding the teaching of their children.

development of a signifier set for soft metadata. Van Garderen working with the Cynefin model and the SenseMaking Triads in the tradition of Snowden [6], I attending with our⁶ Knowledge in Formation model (KiF-model) [7] and Richmond's Trikons [8] in my mind, see section 3 for differences between both approaches.

After fifty stories it proved possible to sketch the four most relevant processes in a KiF-model and to develop a trikonic signifier set for the focal process. In 2013 I had the opportunity to follow two cases in detail. This offered the possibility to collect the scores of all stakeholders involved in one and the same model. In the first case it proved not possible to organize a meeting of all stakeholders. The negotiation process between parent and school remained a zero-sum game. Father and son recently voted with their feet in order to avoid a second criminal prosecution procedure. In the second case it proved possible to organize two such meetings and a win-win solution is realized:⁷ After more than three years this son receives education again. It is worthwhile to note that the same solution would also be agreeable to the parent and child of the first case. In 2.1 I briefly describe the initial narrative environment and the hard metadata. In 2.2 I present the educational KiF-schema and the attached soft metadata.

This is not the place to delve into the content of the stories gathered in order to find some patterns. The interested reader, who mastered Dutch is referred to *De Dunne lijn; tussen naar school en thuiszitten* (The thin line between going to school and sitting at home), *De dunne lijn; perspectieven verbinden* (The thin line, connecting perspectives).⁸ And, for one of the case descriptions *Terug naar af Wouter!* (Back to the start Wouter!).⁹

2.1 Stories and hard metadata

The visitor of our narrative environment first got a short introduction in the subject of interest and our goals. In the narration environment the first, and only obligatory, question we asked is: Tell your story. There are no restrictions to length in the accompanying text field. The idea is that people who go to the internet and tell their story because they experience(d) some frustrating problems are inclined to concentrate on what they deem most relevant for the state they find themselves in. Immediately after this question we asked for a title and keywords. This can be regarded as inviting user generated metadata that suffer heavily from inconsistencies, as is pointed out by Kalbach [2]. This however was not the intention. Primarily we wanted the respondent to take a somewhat distanced look at the story and at the same time prepare the respondent for a final open question: Do you have suggestions based on your experiences? Secondly we tried to enhance the number of clues for judging the measure of credibility of the story told. The invitation to provide suggestions was put after the hard metadata questions that, besides being questions, may set off some

⁶ Sarbo, Farkas and Van Breemen.

⁷ I don't claim that the domain specific KiF-model forces a win-win strategy. At the most it invites one by sorting out the diverse processes and arguments.

⁸ Retrievable from <http://www.koesch.nl/onderzoek/>.

⁹ Retrievable from <http://www.koesch.nl/wordpress/wp-content/uploads/casuswouterd11.pdf>.

association patterns, related to important facts of the problem area in the mind of the respondent.

Next we asked for facts (In order to give the reader at least some insight in the domain, responses are provided after a number of closed questions. N=58):

- What was the duration of the absence?
Some days 9%, some weeks 10%, two months < 16%, 6 months < 49%, was never absent 19%.
- What was the age of the pupil?
The main gain of this question is that by now we know we have to append the question 'duration of absence' with '1 year' and '2 years'. We got answers like from 5 to 10 years of age and 12 to 16 years of age. In general problems surface around 5 (entering primary school), around 10 (entering second part primary school), around 12 (entering secondary school) and around 15 years of age.
- Did the pupil attend a school for special education?
Yes 52%, no 48% (special education takes about 10% of all pupils).
- What kind of school did the pupil attend?
Kindergarten 12%, primary education 40%, secondary professional basic (vmbo) 27% professional follow up (mbo) 19%, general theoretical (havo) 8%, preparatory for university (vwo) 6%. Our evidence shows the occurrence of a downgrading of level, before a drop out.
- I am
Pupil 3%, parent 79 %, grandparent 2%, friend 3%, teacher 2%, ambulant begeleider (mentors of the extra funding for pupils in regular education) 9%.
- (Open question) What were the consequences of the experiences?
- How long did those consequences last?
Some weeks 9%, 3 months 5%, 6 months 9%, a year or longer 64%, there were no consequences 0%.
- Does the pupil attend school currently?
Yes 53%, No 47%.
- If there has been or is a drop out, has it been reported to the officials?
Yes 58%, no 36 %, not applicable 5%, I don't know 12%.
- Do you have suggestions based on your experiences? (lots of suggestions)
- May we know who you are, if so please provide contact details?
7 respondents did not leave contact details.

The basic idea behind the set up thus far is that with the help of (a selection of) the hard metadata and eventual some text search (for instance autism and related terms) in the stories, a user of the database can retrieve all stories that fulfill the conditions. Thus (s)he is enabled to consult the stories or scenario's and decide which scripts may be promising and which scripts aren't for the problem (s)he is facing.

2.2 The KiF-model

This is not the place to go into the depths of the KiF-model regarded as a model that provides a general description of interpretation processes as a procedure. In Fig. 1 the

processing schema is given. The interested reader may consult [7] for details. Here it is enough to stick to the main line.

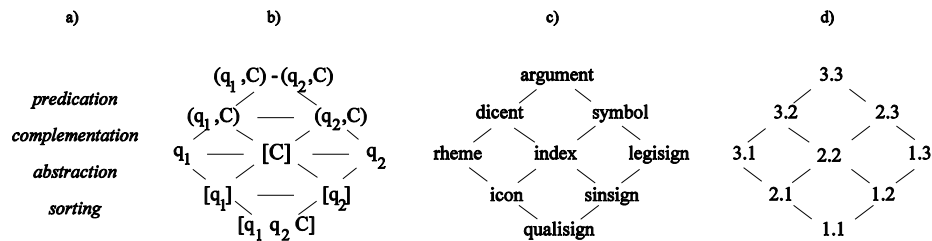


Fig. 1. a) the four stages that can be discerned in interpretation processes, b) the KiF process model of Sarbo and Farkas with the interaction of two variables q_1 and q_2 and the context C , c) the Peircean semiotic terms and d) the Peircean categorical values of every position on the corresponding positions.

If in analyze mode we look at an interpretation process bottom up,¹⁰ we set off with an interpreting system in a certain *state* q_1 that experiences an *effect* q_2 ,¹¹ but at this point both are indiscriminately present, which is represented by the square brackets in 1.b. Since the very same effect may in different interpretation processes yield different *results* $(q_1, C) - (q_2, C)$, the C represents the context and the goals, operative in the interpretive system, that determine the inclination towards the result that is going to be worked with in the stages to come.¹² From a top down perspective: at this point nothing follows, so q_1 and q_2 are both still regarded false, which can be regarded as an expression of doubt, see figure 2 bottom.

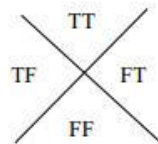


Fig.2. The truth values of two variables distributed over the four compartments of the KiF-model.¹³

First the state and effect are sorted out, but just that, next they must be abstracted, or ‘recognized’ (expressed in terms of fig. 1 b) $[q_1$ (in the ‘negation’ of q_2) yields $[q_1]$ and $[q_2$ (in the ‘negation’ of q_1) yields $[q_2]$).¹⁴ In terms of Fig. 2 the

¹⁰ This does not mean we defend the bottom up stance in the old discussion about reading (bottom up vs top down). It is a mixed affair and some people are more inclined to the one or the other strategy. A person overdoing constantly in one of the directions will end up in a niche in the educational system.

¹¹ Notice that this is in line with Stampers actualism.

¹² What is taken as input and what as noise in the receptive field depends on C , later on we will see that also emotional predispositions bear on the outcome of interpretation processes.

¹³ This is an application of Peirce’s X-box which he used to derive the 16 Boolean operators.

¹⁴ This triadic relation underlies the dependency structure of the KiF-model, cf. $q_1, [C] \rightarrow (q_1, C)$. This also is the relation between State, Effect \rightarrow Response, that summarizes the lot. The KiF-model aims at a description of the argument in semiotic terms, cf 1c.

interpreting system must manage to make the state true for itself independent from but in light of (or in the universe co-determined by) the effect and the effect independent from but in light of the state. When ‘recognized’, to this end the memory [C] is addressed by q1 and q2 for complementation in the light of the goal that is operative. If as well q1 as q2 can be complemented by C - expressed by (q1,C), (q2,C) - and (q1,C) can be unified with (q2,C) (predication stage), a response can be given to the original input. Put in terms of Fig. 2: the original doubt (FF) proved to be resolvable into belief (TT). This of course does not mean that the belief is true; it only means that *this* interpreting system managed to arrive at a result on *this* input.

2.3 The KiF-model in the educational domain

Learning problems experienced by pupils in class, which cannot be met in class, enter school level. Thus we have a process that has as its goal the learning activity of a pupil: Pupil(S), Class(E) ->Learning(R). If this process fails because either the S or the E cannot be made true to a satisfactory degree or both cannot be unified, a new process sets off - Pupil, School->Learning result - of which the former process is a sub-process, see Fig. 3a). If the discussion arrives at this level a serious question is regularly raised: is the class responsible for the failure or is it due to the parents? See for instance Stoutjesdijk [10], but also our stories give ample evidence for this move. So, as a parallel process we have Pupil, Parents->Interest in school, Fig. 3b)

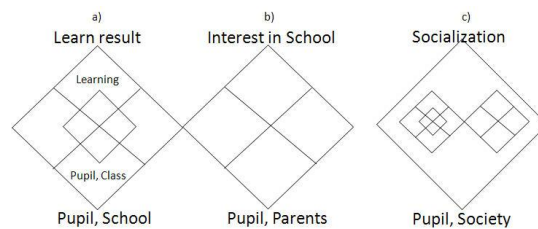


Fig. 3. After a pupil is inscribed at a school process, it enters the class sub processes. These processes are defined by their goal a). If problems arise that cannot be met in class, the discussion is transferred to the super process, but then a parallel process may be deemed relevant and questions arise about the contribution of the parent to the problem b). Both processes have society as their root. Besides that auxiliary processes may be started, indicated by the empty space c). The goal of the root process ought to pose restrictions on what can be argued about the sub processes.

Via the pupil the three processes are related and a variety of subsidiary processes may arise. Being tested by a psychologist¹⁵ and being prosecuted by law¹⁶ are only two of them. In Fig. 3c) Society is the agent that has its Effect on the State of the Pupil. The perspective from which this interaction must be looked at, is its goal ‘Socialization’. This root process has two sub-processes of which one has a sub-process. Besides that, other parallel processes might be involved, indicated by the empty space. We must not assume the law of excluded middle holds, in our attempt to find a solution for a pupil.

¹⁵ Psychologist (S), Pupil (E) -> Testreport (R) or Pupil(S), Psychologist(E)->Remediation(R).

¹⁶ Pupil (S), Public servant (E) -> Convicted (R).

In this exposition the focus will be on the process Pupil, School->Learn result. We do not want to model this process; we want to be able to inspect it with regard to the question what the hindrances are for the achievement of a satisfactorily learning result. Since the pupil will have to do the interpretation, needed for the kernel process *learning*, we want to know whether there is a (negative) bias in the propensity to sort, abstract and complement. Next we want to know which characteristics of the child or/and of the school contribute to the problem. Finally we want to know something about propensities in the predication or unification stage. Cf. Fig.s 1a) and 5.

2.4 Soft metadata and Trikons

Richmond [2005], following Peirce, introduces trichotomy as “the art of making threefold distinctions”. “this new applied science”, he continues, “rests upon Peirce’s triadic categorical distinctions”. Sarbo and Farkas [9] offer a way to illuminate¹⁷ the difference between triadic and trichotomic distinctions. Interpretation processes yield signs as results. On Peircean principles a sign is something that (1) can be regarded in itself, (2) can be regarded as to its relation with its object and (3) can be regarded in the way in which it addresses its interpretant thought. If one of those relations is missing we do not have something that functions as a sign. In terms of Fig. 4 a) out of each of the bold lines a sign takes one position with the restriction that the position on a higher category cannot have a higher value than the position occupied in lower categories. As a consequence ten different sign types can be distinguished. So, in terms of Fig. 1c), a Rhematic, iconic legisign is a sign type, a symbolic, dicent legisign another. This means that for each of the triadic sign relations there is a rule restricted choice out of three, again categorically ordered, possibilities, the 1, 2 and 3 on each bold line. These are each trichotomic distinctions and not triadic relations.

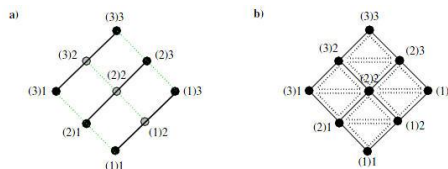


Fig. 4. Each bold line in a) presents a trichotomic distinction pertaining to signs. We get a triadic sign relation by picking a term out of each of the three trichotomies. There is a restriction: the term of a lower trichotomy may not be of lower value than the term picked in a higher trichotomy. On account of this restriction, the Trikons can be ordered in the KiF-model: Trikon (1) in the bottom square, (2) in the square left and (3) in the top square, see Fig. b). The square right is reserved for the effect processed by the interpreting system.

So, what do we want to know about the Pupil? In the first Trikon we want to know something about the sign in itself, the Pupil in this case. Especially we want to know its inclination in the sorting, abstraction and complementation phases. A shy withdrawn or fearful attitude (1)1 leads to different response patterns than an angry or

¹⁷ A full account would demand a treatment of Peirce’s extended sign system and of the differences between the static figure of semiotic terms, 1c), and the actualist, dynamical KiF-model 1b). This falls far out of the scope of this paper.

too assertive externally oriented inclination (1)2, which again differs in its results from a reasoned response that takes the situation into account (1)3.

In the second Trikon we want to know about the Pupils relation with its objects. Is it just the interaction with informational content that is going to make school difficult? (2)1. Is it Pupils interaction with the physical world? (2)2. Or, is it Pupils interaction with other individuals that also process information and interact with the physical world? (2)3.

Finally, we want to know how the unification process runs in the interactions between Pupil (and Pupils caretakers) and School. Are there plans without execution, but nobody makes a point of it? (3)1. Is there hostile disagreement about the action that ought to be taken? (3)2 or are all working towards a solution? (3)3. See Fig. 5.

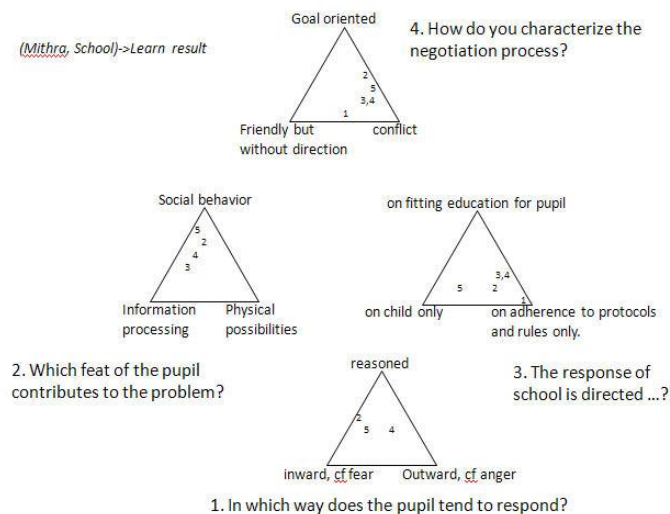


Fig. 5. The four Trikons in their KiF arrangement. The scores are from case Mithra (unpublished) in which I utilized the model to guide negotiation. 1 is the mother, 2 a teacher, 3 and 4 are pedagogical specialists. 5 is a civil servant of child care.

If we were interested in the way School interacts with Pupils in general, then we had to make a KiF-instance with school in the state position. But we are not, we only want to know how School (E) deals with this niche Pupil. This belongs on categorical level (2). School may order all kinds of tests, but leave the problem as it is in the everyday affairs (2)1, it may have protocols that are applied blindly (2)2 or it may decide its course in interaction with Pupil (and Pupils caretakers) (2)3. See Fig. 5.

Since the points in the Trikons can be calculated, it is possible to use them as pointers to stories. And, they do what Kalbach asked for. They do, together with the kiF-model, put data in a meaningful information architecture. The combination is even able to direct the negotiation process in an effective way. On top of that the educational model can be extended with embedded and parallel processes dressed with Trikons.

There is a serious objection that can be raised. It proves to be the case that even in one case different roles diverge in the scores they make. So, what is the value of a story base? Well, yes anyone taking the stories told uncritically for truth is ill advised.

But, as you see in the scores in Fig. 5, the score on the top Trikon already shows that opinions strongly diverge and that it is up to the reader to sort out the promising scripts. And, even if all stories could be trusted, it would be ill advised to just do what proved to be successful. For one thing because the people concerned will not be the same, for another because the goals that are set for Pupils by their parents, will differ in terms of ambition. But also, the goal of this system is not to automate responses to problems that do arise, the goal is to empower the user by providing a model that systematically decomposes the problem into a set of questions and to deliver scripts and scenario's that other people met in dealing with similar problems.

3 Related research

The approach proposed here is akin to cultural anthropological field work in complex societies assisted by ICT. To my knowledge, Kurz and Snowden originated this narrative approach [6]. It is quite remarkable how close the resemblance is between the Cynefin model and the KiF-model and between Triads and Trikons. There are however some significant differences. Triads are constructed in sessions with stakeholders and the terms at the corners must be in opposition with each other, Trikons are constructed in relation to goal oriented domain specific KiF-processes and the terms must adhere to categorical rules. This architectonic improves the chance that data islands, tagged with Trikons can be more easily related to each other.

Richmond's Trichotomic [8] and Keidel's Triangulation [12] are very similar. Richmond bases his Trikons on the categorical distinctions of Peirce and he applies them in grand scale philosophical musings, whereas Keidel developed his triangles in organizational science and describes them in organizational terms: autonomy (monadic, first category), hierarchy (dyadic, second category), network (triadic, third category). Anyone interested in the principles behind the Trikons is referred to Richmond, but if the interest goes to the application of Trikons, Keidel has much more to offer, cf. pp 153-167 of [12] for lots of orderly arranged triangles. In its result, the main difference between Keidel and Richmond is that Keidel uses points and lines in his Triangles¹⁸ in order to signify situations or the direction of change in time respectively, whereas Richmond utilizes the outline of the Triads in order to indicate directionality. Technically, Trikons as I use them are ternary plots.

An important difference between Keidel and Richmond on the one hand and the proposed method on the other is that the former builds Trikonic or Triangular edifices whereas the latter integrates the Trikons with an information processing model.

4 Conclusions

The proposed narrative approach, with domain specific Kif-models that ensure a common frame of mind and deliver meaningful information architectures at the same time, and with Trikons, in order to map and retrieve perceptions, can become a sound

¹⁸ Keidel is the first I know of who did.

method for dealing with quality of service in the long tail. One can imagine different ways to proceed. On categorical grounds: 1. Just throw out this conceptual net on the internet and see what the catch will be. 2. Organize for clusters of schools an environment where stories can be delivered by parents or teachers and let the schools and parents deal with it themselves. 3. Organize a specialized group of civil servants that operates on a regional or national scale and give them the task to solve problems that are brought to the fore and to follow the results in order to improve performance in the long run. In the Netherlands the local civil servants that guard the obligation to attend school and/or the national educational consultants are the most obvious candidates, but only if experienced teachers are added to the group: as it is now, the role of the servants is far too formal and their knowledge of education too scant to be able to deal with the content.

A yet unanswered, but important research question is whether it is possible to make models on the same principles in other domains, as for instance health care and child care, and in other cultures where the goals differ. The movie 一个都不能少 (Not one less), by Zang Yimou, clearly shows that in a metropolis as Shanghai it may be feasible to shift attention from the head to the tail, but that in predominantly rural areas first things still have to come first. Although, ... a lot can be learned from the devotion with which the main character handles the exception she meets.

References

1. Anderson, C.: Makers: the new industrial revolution. Crown business books, Random House Inc., New York (2012)
2. Kalbach, J.: Navigating the Long Tail. In: Bulletin of the American Society for Information Science and Technology, Vol. 34, Number 2 (2008) 36-38
3. Geertz, C.: Thick description. In: The interpretation of cultures, Basic Books, Inc. (1973) 3-33
4. Ginzburg, C.: Clues: Morelli, Freud and Sherlock Holmes. In: The sign of three, ed. Eco, U, Seebock, T.A.: Indiana university press, Bloomington. (1988) 81-119
5. Schank,R.: Virtual learning; a revolutionary approach to building a highly skilled workforce. McGraw-Hill (1997)
6. Kurtz C.F., Snowden D.J.: The new dynamics of strategy: Sense-making in a complex and complicated World. *IBM Systems Journal*, Vol. 42, 3 (2003) 462 – 483
7. Sarbo, J., Farkas, J., & Breemen, A. J.J. van: Knowledge in Formation: A computational theory of interpretation. Berlin: Springer. (2011)
8. Richmond, G.: Outline of trikonik▷* k: Diagrammatic Trichotomic. In: Conceptual Structures: Common Semantics for Sharing Knowledge. Lecture Notes in Computer Science, Volume 3596 (2005) 453-466
9. Sarbo, J. Farkas, J.: Towards Meaningful Information Processing: A unifying representation for Peirce's sign types. *Signs - international Journal of semiotics*, Vol 7 (2013) 1-44
10. Stoutjesdijk, R.: Children with emotional and behavioral disorders in special education: Placement, progress, and family functioning. Doctoral thesis, Leiden University. (2014)
11. Kurz, C.F.: Working with stories in your community or organization. Third edition. Retrieval from: <http://www.workingwithstories.org/>
12. Keidel, R.W.: Seeing Organizational Patterns; a new theory and language of organizational design. Berret-Koehler Publishers, San Francisco (1995)