

# An Ark for the Exaflood Rushing upon Us

Mei Lin Fung<sup>1</sup> and Robert S. Stephenson<sup>2</sup>

<sup>1</sup> Institute for Service Organization Excellence, Palo Alto, California USA,

<sup>2</sup> The Tech Museum, San José, California USA

**Abstract.** As of March 2010, 21 exabytes ( $2.1 \times 10^{19}$  bytes) were being added to the Internet each month – an “exaflood.” Making sense of this data flood is one of our greatest challenges. What role might knowledge federation play (to switch metaphors) in weaving these data into knowledge? We provide a sketch of a distributed system for answering questions posed online that we call the Loom. This is a human system augmented with a bit of technology. At its core is a reputation system to track the skills of the participants and their domain(s) of expertise. Payments made by those whose questions were answered reward and sustain its participants. The roles of its members include 1) those who pose questions, 2) those who gather and frame the questions, 3) those who refer questions to experts they know, 4) the answerers, 5) those responsible for curating questions and answers so they are accessible for reuse. The technical parts of the Loom are a communication system, a repository and a reward currency.

**Keywords:** sensemaking, reputation system, social network, knowledge gardening, knowledge federation

## 1 Introduction

From the beginning of humanity to 1999, a UC Berkeley study estimated that we created 12 exabytes of information. An exabyte is a billion billion bytes, or 50,000 years of DVD quality video. Wikipedia’s entry for Exabyte<sup>3</sup> estimates that Internet traffic reached 21 exabytes per month in March of 2010 – and this rate is doubling every year. Entire categories of jobs are changing as print media vehicles jettison their journalists, schools re-think teachers, and everyone tries to keep afloat as the tidal waves of information break on every human endeavor. The Singapore National Library is organizing networked communities inspired by Douglas Engelbart. Social networks are being explored by the US Department of Defense. The US is brimming with collaborative initiatives like Huffington Post (journalism), DreamFish (jobs), TEDxSalonBayArea (conversation), QuantifiedSelf (health) and new grass roots movements are now emerging that span or cross industries like Government, Climate Change, Energy and more. Broadcast TV is being overtaken by YouTube videos. What role does knowledge federation play in nurturing and supporting these emerging human connections?

<sup>3</sup> <http://en.wikipedia.org/wiki/Exabyte>

Upon this gifted age, in its dark hour  
 Falls from the sky a meteoric shower  
 Of facts.... they lie unquestioned, uncombined  
 Wisdom enough to leech us of our ill  
 Is daily spun; but there exists no loom  
 To weave it into fabric  
 – Edna St. Vincent Millay

Upon our gifted age in its new awakening  
 Falls from the sky a meteoric shower of stories  
 Of acts of beauty;  
 Compassion that lies unremarked and unrecognized  
 Wisdom to nurture life, heal hurts, grow joy, is daily spun  
 Let us commit our loom of civilized intention  
 to weave wisdom into the fabric of our souls  
 – Mei Lin Fung

There is no loom except our own minds: technology can fetch and correlate but it cannot understand or make sense of this shower of facts – at least not now or anytime soon. The issue becomes one of finding and consulting the few who have the best understanding of that which we seek to “weave.” This is actually two problems: 1) what do we mean by “best” and 2) how do we find those who are the best in that domain. Answering the former and succeeding at the latter requires a system for reputation and trust. Humans have been dealing with each other based on reputation for as long as we have been communicating, and we have developed very subtle systems of reputation and trust that work best in a face-to-face environment, or at least in a restricted group where many people know each other.

The system we propose, the Loom, is intended rather as a distributed system that could work over the Internet among people who are mostly strangers to one another. It is primarily a human system, augmented with a bit of technology. It is, in other words, a “Mechanical Turk.”<sup>4</sup> The Loom is a tool to create meaning and sense out of the exaflood. In doing so, it will build our Ark against the flood.

## 2 The Loom Reputation System

There are many examples of distributed, computerized reputation systems – for example credit reporting agencies, Amazon’s book recommendation system or SlashDot’s comment rating system – but they are generally one-dimensional. They treat reputation as a single quantity, whereas it is, in fact, a vector of high dimensionality. For example, I cannot think of anyone I would recommend more for advice on the Java programming language than James Gosling, but if the issue were selecting a suit I would probably prefer to ask George Clooney.

<sup>4</sup> A celebrated chess-playing machine of the 19th Century, operated in fact by a small man hidden inside. See [http://en.wikipedia.org/wiki/The\\_Turk](http://en.wikipedia.org/wiki/The_Turk)

We start from the premise that there are as many dimensions of reputation as there are domains of human knowledge where someone might reasonably have a question that needs answering.

The ingredients of the Loom reputation system are eight:

1. people with important (to them) questions to ask online (“questioners”),
2. people who identify interesting questions “in the wild” and submit them to the Loom (“framers”),
3. a system of communication,
4. a network of people who know stuff (“answerers”),
5. people who know other people who might know stuff (“referrers”),
6. a record-keeping system,
7. the people who tend to it (“gardeners”), and
8. a form of reward or currency.

### 2.1 People with Questions

These may be anyone who cares especially to get an answer, enough to pay for it. The answers may require specialized knowledge, particular wisdom or just common sense. Answers will generally not exceed one page of text. The questions, too, should be expressed in a page or less. Of course, many important questions cannot be posed or answered in the compass of a single page. For these, the Loom can only serve as a referral system: “can someone recommend an architect to remodel a modest private home in Berkeley, California?”

### 2.2 People who Frame Questions

The framers provide a human interface to the Loom, when it’s needed. Framers seek out those with pressing questions, help to screen and reframe their queries and act as a buffer between them and the Loom’s systems. They are interpreters and salespeople, and will play an important role in determining the Loom’s success.

### 2.3 Communication System

By default, all communications will be anonymous. Although this may sound surprising for a system of reputation, it is in general a necessary feature. Many knowledgeable people are also busy, and may be reluctant to give out their personal contact information for fear of receiving unwanted contacts. Finally, this helps ensure that answers are judged on their merits, rather than the answerer’s authority. Of course, parties may choose to add their names if they so choose; anonymity is only a default. Most communications will pass through a central server to preserve anonymity. All messages exchanged, though anonymous, are posted publicly. To reduce latency, it is important that the server support many modes of communication – SMS, phone, IM, email, Twitter, Facebook, etc. – so that each member of the network can interact with it in the manner she finds most convenient.

## 2.4 Network of Referrers and Answerers

Each of these people has a profile in the Loom reputation system. The initial group used to bootstrap this system will be the group of KF10 attendees.

## 2.5 Record-Keeping System

This is the common server that forms the hub of the system, assures anonymity, tracks all interchanges, and maintains the history and reputations of all the Loom’s participants. Most importantly, it is the repository of all questions and their answers.

## 2.6 Knowledge Gardeners

The gardeners curate these questions and answers, so they can be as useful as possible in the future. They format the answers so they may be found by relevant queries, and qualify them so that their limitations and domain of applicability are also clear. They are what Douglas Engelbart referred to as “certified public logicians.”<sup>5</sup>

## 2.7 Currency

The currency is used to reward successful answerers and referrers. Initially it may be a virtual currency, but for true scalability it will need to be interchangeable with conventional currencies. To provide anonymity, a currency like Bitcoin<sup>6</sup> might be appropriate.

# 3 How the Loom Weaves

Consider the following scenario:

## 3.1 The Question

1. Quentin is curious about compostable toilets but does not know about the Loom. His friend Fred does, however, and enters Quentin’s question as follows: “is there a design for a compostable toilet that makes sense for a house in the suburbs?” Fred is the “framer.”

---

<sup>5</sup> <http://engelbartbookdialogues.wordpress.com/help-us-raise-collective-intelligence/engelbart-on-structured-arguments/>

<sup>6</sup> A peer-to-peer digital currency, see <http://bitcoin.org>

### 3.2 Referral

2. The system checks its registry to see if any similar question had been answered before. None had, so Quentin's question is routed to several people, including Rita, who have good reputations as referrers on a broad range of issues.
3. Rita refers the question to Robert, since she knows he is interested in sustainable architecture.
4. Robert forwards the question to Agnes, because he knows she has just finished building a sustainable house.

### 3.3 Negotiating

5. Agnes replies that the question is incomplete: it needs to specify how much the toilet builder is willing to pay for installation and maintenance. After a quick consultation with Quentin, Fred adds these details to his posted question.
6. Quentin and Agnes negotiate a price for the answer. Quentin offers 10 BTC (Bitcoins) for a complete answer and Agnes accepts.
7. Quentin receives several other offers through Fred. He chooses Agnes's, based on Robert's recommendation, and puts the 10 BTC fee in escrow.

### 3.4 Answering

8. Agnes forwards her answer to the question, via Fred, to Quentin.

### 3.5 Settling

9. Quentin reviews Agnes's answer and, after consideration, decides that, although it did not completely fit his situation, Agnes's answer satisfied 80% of his question. Agnes agrees to 80% and is satisfied. Two BTC of the fee is returned to Quentin and the remaining 8 BTC will be released from escrow as described below.

### 3.6 Integration

10. To ensure that Agnes's answer be as useful as possible in future, George integrates and cross-references it into the repository so it can be found again. He indicates the limitations of the answer, making clear, for example, where it fell short of Quentin's expectations.

### 3.7 Payment

11. Rita and Robert both receive 5% of that for their referrals. Fred receives 2% for submitting Quentin's question and helping to refine it. George receives 2% for curating its answer. The Loom system fee is 1%. Agnes receives 85%, or 6.8 BTC, for providing the answer.

### 3.8 Remembering

12. The system preserves the following information:
- Quentin and Robert:** the question, that it had to be clarified, that Quentin was 80% satisfied; how much was paid.
  - Fred and Rita:** the question, that Agnes bid on it, that her answer was 80% satisfactory. That Agnes agreed that was fair.
  - Agnes:** the question, the clarification, the price, the answer, that it was 80% satisfactory, that she was satisfied.

Should another questioner ask the same question, she would immediately receive Agnes's bid at 50% of the original price. Should she ask a different but related question, or balk at the fee, her question would be relayed directly to Agnes for her to bid on. In either case, if Agnes's bid was accepted and her answer satisfactory, the original referrers (Rita and Robert) would each receive 4%; Fred and George 1% each; the Loom system 1% and Agnes 89%. As Agnes answered subsequent questions related to the first one, the other participants' shares would decline by a percent (e.g. 5%, 4%, 3%, ...) until they became zero. The Loom system always gets 1%. In some cases, questions might be answered by a team of individuals who would share the answerer's reward.

The loom network grows by invitation, at least beyond the walls of the KF10 conference. Someone in the network knows someone who would be a great resource in a specific area, or would make a great referrer. As each new person joins, her reputation would initially be unknown, but will grow over time in measure with her contributions.

## 4 Challenge

This blueprint is obviously incomplete and only a sketch. It lacks essential details about the central server and the reputation system itself. The Loom network has yet to be built. We gift it to KF10, for whatever it's worth. We challenge you, the KF10 participants, to take it, make it your own, fill it out and warp it to your needs. If you think it has promise, our dream is that we, the conference attendees – real and virtual – may start using it in the domain of, let us say, knowledge representation/storage/retrieval/federation.

**Acknowledgments.** We, the authors, are grateful to all the members of the Knowledge Federation, Continuing the Conversation and Program for the Future communities for their feedback and support. We would like to recognize Mr. Jack Park in particular for his many thoughtful comments that have influenced the direction of this paper.