

Socrates, Trust and the Internet

Kieron O'Hara

Dept of Electronics and Computer Science
University of Southampton
Highfield
Southampton SO17 1BJ
United Kingdom

Abstract: Socrates, one of the world's greatest philosophers, never wrote anything, and confined all his philosophy to spoken debate. The important issues for Socrates were trust and control: he felt the radical decontextualisation that resulted from the portability and stasis of written forms would obscure the author's intentions, and allow the misuse of the written outside of the local context. Trust has once more become a central problem, both politically and epistemologically, but since Socrates' day, various technologies have undermined his distinction, making the relationship between trustworthiness and linguistic mode more complex. In this paper, I review the state of the art in Internet technologies, showing (a) how developers and authors attempt to establish trust in their websites or e-commerce processes, and (b) how new work in dynamic content creation further blurs the spoken/written and global/local distinctions.

In his early work the *Phaedrus*, Plato represents Socrates as arguing as follows.

SOCRATES: You know, Phaedrus, writing shares a strange feature with painting. The offspring of painting stand there as if they are alive, but if anyone asks them anything, they remain most solemnly silent. The same is true of written words. You'd think they were speaking as if they had some understanding, but if you question anything that has been said because you want to learn more, it continues to signify just that very same thing forever. When it has once been written down, every discourse roams about everywhere, reaching indiscriminately those with understanding no less than those who have no business with it, and it doesn't know to whom it should speak and to whom it should not. And when it is faulted and attacked unfairly, it always needs its father's support; alone, it can neither defend itself nor come to its own support. (Plato 1997, 275de, p.552)

This argument that writing is untrustworthy, compared to speech, has four components. Firstly, the writing is not open to question, discussion, or interactive scrutiny. Second, the author has no control over who reads a piece of writing once it is out of his hands. Third, there are ideas, facts or thoughts that should not reach certain groups of people. Fourth, the author of a work is not in general available to clarify the points it makes. These boil down to the claim that, for a passage to be trusted, the interlocutor must be able to secure a personal acquaintance with the originator of the passage. The words, unadorned, cannot be trusted by themselves, without an interrogation, an evaluation process, which typically will involve the presence of the author (and in most circumstances, a spoken interchange).

This type of trust, based on personal acquaintance, is explicitly *local*. What does such trust buy the truster? Well, trust is used in conditions of uncertainty, typically but not always about the future. Trust involves assuming true without further investigation the *bona fides* of another actor, process or institution. In such circumstances, the information-processing costs for the truster are reduced, and her planning is made correspondingly easier; the risk she runs, of course, is that the trustee is not reliable, and

is making false claims. If the trust is betrayed, then the truster's plans will not conclude successfully. When trust is local, that means that the truster accepts the *bona fides* of people she already knows, and therefore either already has a pretty shrewd idea that their *bona fides* are accurate, or is sure that she can catch up with any miscreants to extract compensation.

There are plenty of societies where trust can remain local; the small Greek city-states such as Socrates' Athens are obvious examples in this context, and consider also the contested election for the Master of a Cambridge College in C.P. Snow's novel *The Masters*, where all the key electioneering is conducted through the medium of the spoken, and the written plays a minor and often obstructive role. However, such a model of trust can be quite inhibiting for more complex or distributed societies, as argued for example in (Fukuyama 1995), which claims that the lack of social structures able to support the trust of non-kin is a key explanatory variable for the relatively poor economic performance of countries such as Italy and China, compared to Japan or the USA.

More complex societies require a transition from local to *global* trust, to a situation where actors are able to trust people or processes of which they have no personal acquaintance whatsoever (O'Hara forthcoming, Chapter Four). One way of promoting global trust is the use of *institutions* for investigation, regulation and sanction of particular types of behaviour. Local trust is still required, because the truster requires some sort of personal acquaintance trust of the trust-supporting institutions themselves (on pain of infinite regress). Nevertheless, such acquaintance is much more efficient, as the act of trusting an institution may lead, by transitivity of the relevant trust relations, to the actor's implicitly trusting all that institution's clients. Equally, systemic risk is also increased, because the default of one of those institution's clients may lead to a failure in trust of the institution, and then the loss of trust of all the clients.

The analogy between the spoken and the written, and the local and the global is suggestive; in fact, *pace* Socrates, the written word is often trusted, and quite properly, with the help of various institutions, such as peer review, or the publicity and certification of an author's provenance (hence my work address at the head of this paper). Institutions such as this carry many of the information-processing costs of examining and regulating potential trustees' *bona fides*. Trust-bearing institutions have often been seen as essential for the creation of truly modern societies (Fukuyama 1995, Misztal 1996).

This makes the case of the Internet, the postmodern space *par excellence*, extremely surprising. The net has a very sparse institutional structure, and in particular, trust regulation is negligible; this despite many worries about security of property on the net (particularly with respect to e-commerce), and about quality of information (both for data sharing between programs and for ordinary sources of knowledge such as academic papers). Hence, despite the global nature of the space – in theory geography is irrelevant to it – trust on the net is often *local* (O'Hara forthcoming, Chapter Five).

The Internet being defined by its architecture, institutional trust could always be built into it, with online certification techniques probably being mature enough to accompany

most interactions (Lessig 1999); indeed, this is happening to some extent, as e-commerce begins to develop, and we see closer integration between banking, credit and information systems. The problem for security on the net is that identities are very fluid; users can adopt sets of properties that (a) need not accurately describe them, and (b) they can drop very easily. Hence ensuring that a retailer will deliver the goods you ordered, that a software download will perform the tasks as specified, that a paper you have downloaded really is by a professor at Harvard, or that a teenager talking to your child in a chatroom really is a teenager, are all non-trivial. On the other hand, it is the possibility of masquerade that is part of the Internet's attraction for many users.

The development of the Internet is a political problem, albeit one that will probably be solved via big commercial pressures. However, the only way global trust will be achieved on the net is via the fixing of identities and the creation of certification institutions, with the concomitant loss of privacy. As it stands, the trust mechanisms available on the net are purely local, such as branding (in e-commerce, such names as Amazon or eBay, or for research purposes, the '.ac.uk' or '.edu' suffixes for example), reputation, or genuine acquaintance. Indeed, in my own field of knowledge acquisition and technologies (a relatively small research field), I find my own trust practices very strongly related to the actual people I have met, conference presentations I have seen, etc – a very powerful personal acquaintance bias. Initiatives such as the web of trust attempt to harness this (Golbeck et al 2003).

However, another approach to trust can trade off Socrates' insight about the advantages of speech. Although the Internet is primarily a written medium, this is partly as a result of the paper-based metaphor used by most interfaces; actually the World Wide Web portion of the net supports many media. These multimedia formats, combined with intelligent context-sensitive processing, allow the written forms of the web to regain some aspects of the spoken, to blur the spoken/written distinction still further.

Let me take for an example a system with the terrible name of Artequakt (Kim et al 2002). This system is designed to produce biographies of artists; a typical scenario of use would be that you are in a museum, you encounter a painting that you like, and you call up Artequakt on your hand-held computer. It makes a wireless connection with the Internet, and generates a biography of the artist, whose name you have typed in. The biography is generated as required; it is not recalled from a database but instead created anew from available documents on the web.

Artequakt exploits a number of technologies. Web *searching* calls up pages relevant to the artist. Because of the relatively circumscribed domain – of artists – many of the common relationships, concepts and objects used in art history discourse can be captured in structured *ontologies*. *Information extraction* techniques are then used to extract significant concepts from natural language accounts of the artist, and to map those concepts onto the ontology. This creates a knowledge base about the artist, what Bal calls a *fabula*, an underlying set of raw facts and chronological relations between events (Bal 1998). Given the restricted nature of the domain, it is possible to generate a smallish number of narrative templates, that can be populated from the fabula; hence context-dependent narratives about the artist's life can be created as and when required.

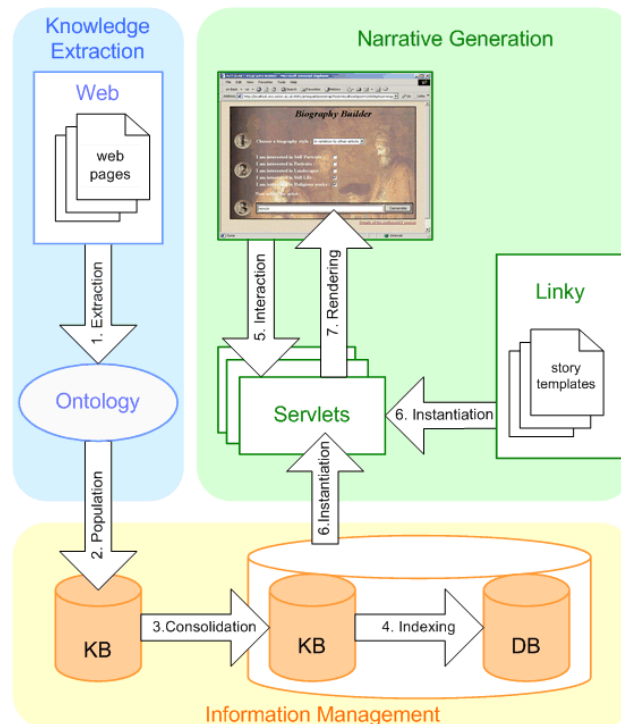


Figure 1: The Artequakt Architecture (from <http://www.aktors.org/technologies/artequakt/>)

The intriguing aspect is that many of the characteristics of the spoken, as set out by Socrates as allowing trustworthiness in the *Phaedrus*, are enabled by this architecture. Narratives are presented to users according to their requirements; it is possible to question the system (though not in the current implementation) to produce further clarification. And indeed, as the state of knowledge about a particular artist changes – and as that change of state is reflected in the content on the web – then the responses of the system will change. In these ways and others, Internet technology is helping blur barriers between spoken and written (as many other technologies have done in the past, down even to Socrates’ day, of course), and thereby present new, non-institutional, local methods of creating and supporting trust.

References

- Bal M. 1998. *Narratology: Introduction to the Theory of Narrative 2nd Edition*. Toronto: University of Toronto Press.
- Fukuyama F. 1995. *Trust: The Social Virtues and the Creation of Prosperity*. New York: Free Press.
- Golbeck J., Hendler J. & Parsia B. 2003. Trust networks on the semantic web. Paper presented at *WWW 2003*, Budapest, Hungary, May 2003.
- Kim S., Alani H., Hall W., Lewis P.H., Millard D.E., Shadbolt N.R. & Weal M.J. 2002. Artequakt: generating tailored biographies with automatically annotated fragments from the web. Paper presented at *Semantic Authoring, Annotation and Knowledge Markup Workshop, ECAI 2002*, Lyon, France, 2002.
- Lessig L. 1995. *Code and Other Laws of Cyberspace*. New York: Basic Books.
- Misztal B. 1996. *Trust in Modern Societies*. Cambridge: Polity Press.
- O’Hara K. forthcoming. *Trust in the Twenty-First Century*. Cambridge: Icon Books.
- Plato. 1997. ‘Phaedrus’ (A. Nehamas & P. Woodruff, trans.) in J.M. Cooper (ed.) *Plato: Complete Works*. Indianapolis: Hackett. 506-556.