Formation Process Tools for creating sustainable Virtual Research Communities

Michael Gardner, Anuroop Shahi and Chris Fowler
Chimera, The University of Essex, Adastral Park, Ipswich, IP5 3RE, UK
mgardner@essex.ac.uk

Introduction

The majority of existing CSCW[1] tools and the underlying framework architectures are based on the assumption that most research communities already exist and are focused on a domain or specific problem (e.g. nuclear physicists, or sending a scientific probe to Mars), and the major challenge is perceived as providing them with tools to allow them to extend from real to the virtual, and to support the variation in domain needs that specialized communities may require. However, there is increasing pressure for research communities to become more multidisciplinary (e.g. the drive for the 'grid'; commercial research as well EU and UK funded Link or multi-disciplinary programmes). Multi-disciplinary teams have special requirements inherent in bringing together often diverse communities and establishing a 'new' unified research community. Whether such communities form effectively and more critically remain stable and sustainable is dependant on the formation process. Such formation processes need to go beyond search facilities for discovering partners (although this is undoubtedly important but such functionality already exist; eg EU's 'Find a partner' service); to addressing the more socio-cultural issues concerned with the formation of virtual research communities much of which are currently not well represented in existing framework architectures.

This paper adopts Wenger's [2] approach to the formation of communities of practice (CoP). The key concept for Wenger is that knowledge is created through social participation. Participation is not simply collaboration (and there are no shortage of collaboration tools – see the CSCW literature [eg 3,4]), and is a 'joint enterprise as understood and continually renegotiated by its members'. Our engagement in tasks is social, involving shared ways of doing things, typified by a discourse with many shared assumptions, short-cuts and rapid flow of information. Finally communities of practice are characterized by a 'shared repertoire of communal resources (eg artifacts, vocabulary, routines etc)'. This sustained mutual relationship which typifies an effective community of practice is our goal, and how to achieve this in a virtual environment is our challenge.

The paper proposes 'new' CoP tools that meet the needs of a specific community of practice. For example we have been studying the CoP centred around a large commercial research organization (BTexact) and its strategic relationship with University centres of research excellence. Many of these more strategic based CoP fail because of the lack of support at their early stages of formation. This is true of the formation of most new CoP's and so we see our work can be generalized. In BT's case the strategic partnership is a triangle of Electrical Engineering excellence (UCL @ Adastral Park); Business excellence (the Judge Institute, University of Cambridge), and Socio-technical excellence (Chimera, University of Essex).

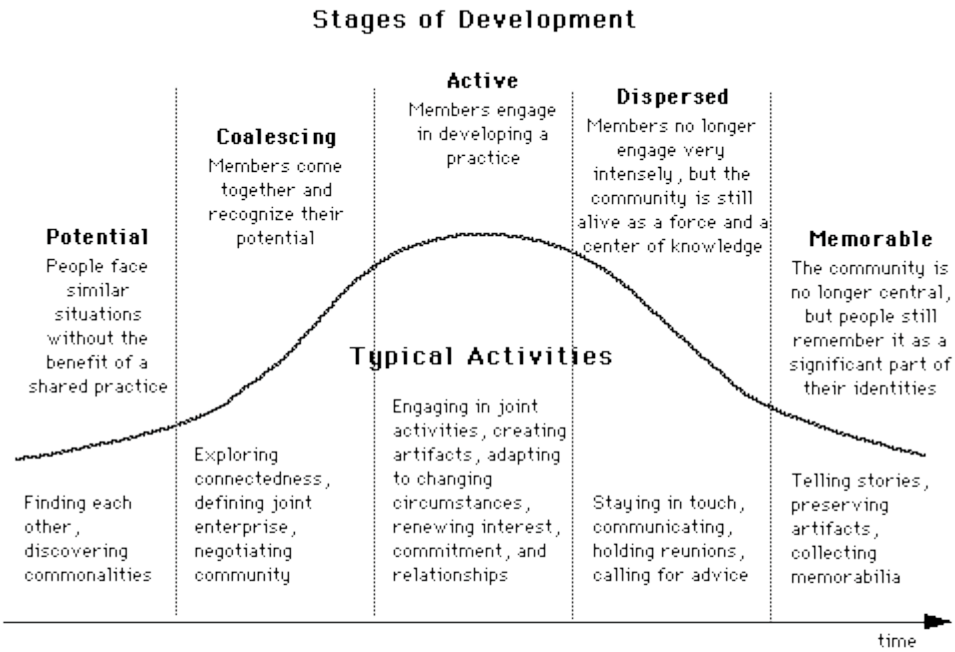
Theoretical Background

The theoretical foundations for deriving the new tools comes from the work of Wenger [2] and Lave and Wenger [5]. For Wenger there are a number of stages in the development of a community of practice (see figure 1 below).

Most tools and most frameworks focus on the 'active' stage only. This paper argues that a whole range of functionality is not currently considered for the first two stages – the formation of the CoP. We propose an architecture, based on existent and emerging technologies/standards, to realize the first two stages of Wenner's approach. This paper also considers how the outcome of both stages can be fed into the 'active' stage thus being used in existing tools and frameworks (for example, see [1]).

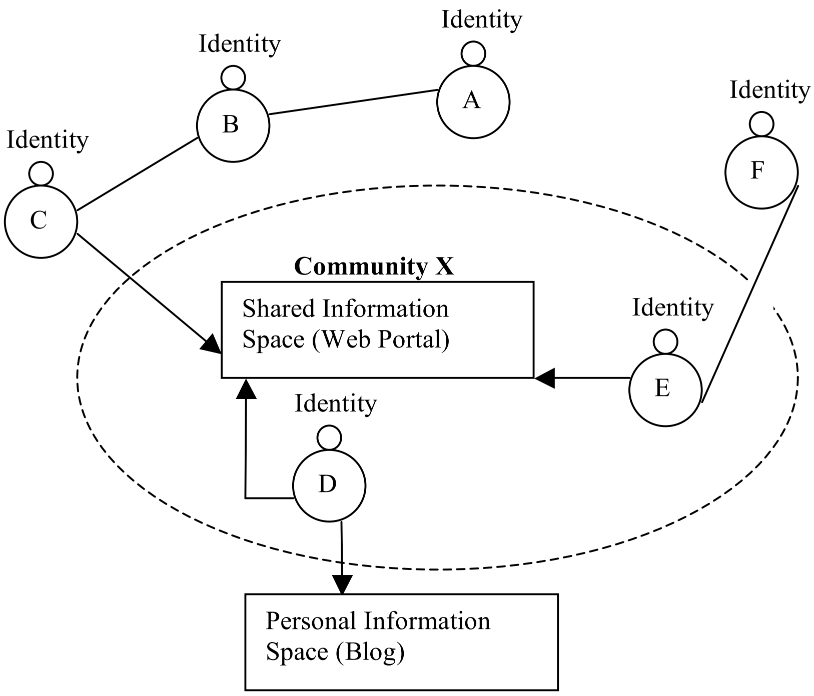
[1] CSCW or Computer Supported Co-operative Work is being used in this report to cover a generic set of tools designed to support collaborative work.

Figure 1. Wenger's community development stages



Donath [6], in her thesis titled 'The design of social environments for electronic communities', proposes a design platform for 'creating' sociable virtual communities. Here, two important design metrics are emphasized: these are the creation of representations of social phenomena (online identity) and the role of information spaces as contexts for communication. Online identity provides a base for evaluating and understanding an entity before conducting an interaction. Information spaces, or online neighborhoods such as community websites, provide communities with a collective memory and cultural vocabulary. Information spaces can be shared, where members of a community update the same space (community web sites, Wikis etc), or personal, where individuals have their own personal spaces, e.g. home pages and blogs etc. Figure 2 illustrates the concept of community formation:

Figure 2. Forming virtual research environments



Using the model (figure 2), we can map the first two stages of Wenger's theory:

The first stage, 'Potential', is based on the discovery of prospective researchers or existing communities. Here, identity plays an important role in 'discovering' researchers. Studying figure 2, researcher A may wish to discover a specific type of researcher, C or Community, X. Researcher A could perform a search by using a large social network of researchers/communities, thus being able to discover C via B. Since researcher C is linked to community A, researchers D and E may also be discovered. Successful discovery will very much rely on an adequate methodology/architecture for identifying 'researchers' and 'research communities'. This could be based on recent software developments in social networks [7].

Wenger's second stage, 'Coalescing', explores connectedness and negotiation to form communities. Connectedness can be explored by examining a relevant 'information space', obtained from the identities of researchers and communities. Information spaces will contain published media and texts, such as blogs, wikis, homepages, portals, papers etc, and will allow the researcher to gain a more extensive and clearer account of connectedness, when compared to studying only the identities of discovered researchers. For example, after discovering community X, researcher C could browse the community's information space, which contains published material, and determine connectedness. Alternatively, members of community A, such as member D, could have their own personal information spaces, such as a personal page or blog. Researcher C could thus link to this information, and determine that researcher D's research is a better match than that of community A, thus choosing to pursue the creation of a 'research relationship' for prospective collaboration with researcher D.

Proposed System

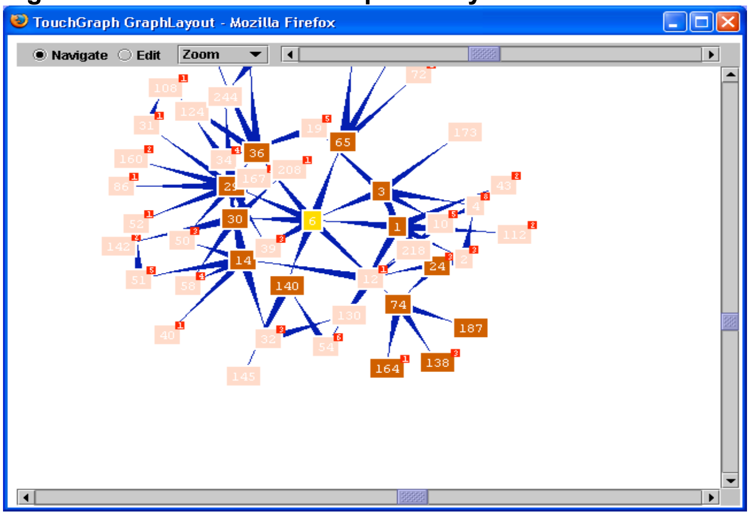
Our intention is to use the friend-of-a-friend vocabulary [8], together with existing research vocabularies, to create a prototype system for the discovery and formation of new research communities. Our approach will be similar to that of the Chandler system [10], where a hybrid architecture is used based on peer-to-peer and client server topologies.

Underlying the system will be the concept of social networks that represent both researchers and existing research communities. Two types of relationships will be emphasized: researcher to researcher and researcher to community Discovery will be possible by traversing these social networks. Figure 3 illustrates a touch graph interface to the system. Users will enter keywords describing the relevant research areas sought, e.g. 'Bluetooth positioning'. The system will then return the relevant results as a clickable social network containing nodes that represent either individual researchers or research communities. Navigating over a node will display a researcher's, or community's, profile together with any links to the relevant information space. Clicking on an individual node will expand the node's social network hence allowing additional nodes to be discovered. We believe that this approach will allow extremely large numbers of researchers and communities to be visualized in a manner where users are not overwhelmed with irrelevant research results, together with allowing users to narrow down discovery by clicking through the various parts of a social network they find relevant.

Technical Architecture

Implementation of the theoretical architecture involves the identification of appropriate technologies to realize the concept of online research identities and information spaces. Ideally, a distributed system based on a peer-to-peer/hybrid architecture for instance, would be most beneficial, with users/institutions having complete control over their research identities/information spaces. Representation of research identities requires a language that is rich enough to express anything in a standardized, extensible and platform independent manner. This language must also support comprehensive search and discovery methods when compared to traditional web based methods. Furthermore, many online communities already exist with existing information spaces such as community web sites, blogs and forums for instance. The system should thus accommodate, and provide mechanisms to seamlessly integrate existing information spaces.

Figure 3. Interface to the Proposed System



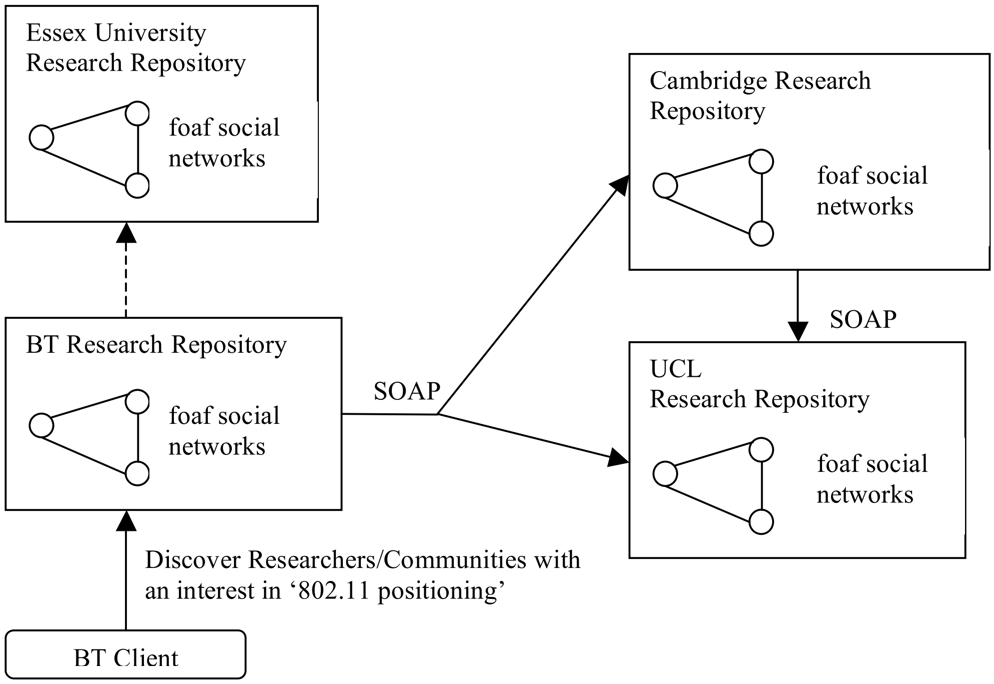
One such profiling protocol is friend-of-a-friend (FOAF), a vocabulary that 'allows machine readable pages for people, groups, companies and other kinds of thing' [8]. Essentially, FOAF provides a useful building block for creating information systems that support online communities [9]. Underlying FOAF is the concept of social network theory, together with open technologies based on the semantic web – some regard FOAF as a semantic web ontology. FOAF profiles can easily be linked into existing semantic web ontologies/vocabularies, since RDF is the underlying knowledge representation language. Information spaces may be linked to FOAF profiles via URLs, due to FOAF adhering to the decentralized architecture of the web.

A sample FOAF profile has been included in Appendix A. The profile is in effect an online identity, which links to existing information spaces such as personal/group homepages, publications and web blogs etc. Using the example, two profiles are presented describing a person and a group.
The person profile includes person specific information such as name, email address, links to personal information spaces: such as web blogs, homepage and publications etc. James's research interests have been described by employing a standard ontology located at http://www.ontologies.org/academicresearch. This ontology could define mappings between various research concepts, e.g. that ubiquitous computing is similar to pervasive computing therefore allowing searches based on semantics rather than pure keywords. In the example, James's social network has been represented using the 'knows' property. As shown, Jackie Brown is one member of James's social network. These social relationships are fundamental in aiding the discovery of researchers when building research communities. To aid people discovery, research communities/groups may also be represented. As shown in the sample FOAF profile, a group has been defined for Essex University Intelligent Buildings. Here, the group has been described detailing its members, such as Jackie Brown, together with an overview of its research interests. Links to the group's information space, such as publications and community homepage are also represented.

Figure 4 illustrates a high-level architecture for the proposed system. Each institution will maintain its own central repository containing FOAF profiles identifying its various researchers and research communities, together with links to relevant information spaces. Each repository will act as a peer in a P2P system, thus being able to serve other 'trusted' repositories. For example, using figure 4, a BT client may wish to discover researchers or communities with interests in '802.11 positioning'. Firstly, the client would send a request to its local repository, which would find researchers/groups within BT. The local BT repository could also act as a client to external repositories, thus issuing discovery requests to both UCL and Cambridge. Additionally, since a P2P approach is taken, new peers may be discovered, such as the research repository present at Essex University.

Integration with Chandler
One interesting scenario would be integration with Chandler, a personal information management system. Chandler could be used to maintain each researcher's FOAF profile. Local repositories could then be updated using Chandler's synchronization features, as mentioned in the Chandler specification for hybrid architectures [10].

Figure 4. Overall Architecture



Contact lists are an integral part of Chandler, as they help provide email, shared calendars and messaging. One problem with FOAF is manual profile maintenance. A person's social network is constantly undergoing change. One way of easing manual maintenance of profiles is to infer relationships from Chandler's contacts component and then implicitly add them to a person's FOAF profile. This way, every time a user adds a contact to the Chandler system, the person's FOAF profile will also be updated to reflect any changes.

Chandler provides an open platform based on the concept of agents. A future extension of this system would be to deploy agents that implicitly organize various Chandler peers into research communities thus monitoring user behavior and dynamically altering a user's social network. Since peers would be ordered into communities/groups, discovery/search the discovery efficiency would be dramatically improved.

For interoperability at a higher level, portlets based on the JSR-168 standard could be incorporated to interface with the system thus acting as a front end service. These portlets could then be integrated into larger Virtual Research Environment (VRE) portals, hence being reused across a wide range of VREs. Chandler's ability to handle web based content will allow for portal integration into Chandler itself, thus allowing the community formation process to be conducted from within Chandler.

References

1. Report of the JCSR VRE Working group, _Roadmap for a virtual research environment_, JISC, 2004.
2. Wenger, E. (1998). _Communities of practice: learning, meaning, and identity_. Cambridge: Cambridge University Press.
3. Marca, D. & Bock, G. (1992) _Groupware: Software for Computer Supported Co-operative Work._ IEEE.
4. Udell, J (1999). _Practical Internet Groupware_. O'Reilly

5. Lave, J., & Wenger, E. (1990). *Situated Learning: Legitimate Periperal Participation.* Cambridge, UK: Cambridge University Press.

6. Donath, J.S (1996) . *Inhabiting the Virtual City: The design of social environments for virtual communities*. MIT http://smg.media.mit.edu/people/judith/Thesis/

7. Scott, J (2000) *Social Network Analysis*, London and Beverley Hills, Sage Publications,

8. Brickley, D & Miller, L. (2004) *FOAF vocabulary specification* http://xmlns.com/FOAF/0.1/

9. Dumbill, E. (2002)., *XML Watch: Finding friends with XML and RDF*. IBM Article for FOAF http://www-106.ibm.com/developerworks/xml/library/x-FOAF.html

10. The Open Source application Foundation. (2003). *Chandler for higher education: requirements and recommendations* http://www.osafoundation.org/Chandler%20for%20Higher%20Education.pdf

11. Van Helvert, J & Fowler, C.J.H. (2004) *Scenario-based User Needs Analysis*. In Ian Alexander and Neil Maiden (eds) *Scenarios & Use Cases Stories through the System Life Cycle.* Wiley: London

APPENDIX A - Sample FOAF Profile

```
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
xmlns:owl="http://www.w3.org/2002/07/owl#" xmlns:FOAF="http://xmlns.com/FOAF/0.1/">
<FOAF:Person>
<FOAF:name>James Brown</FOAF:name>

<FOAF:mbox_sha1sum>241021fb0e6289f92815fc210f9e9137262c252e</FOAF:mbox_sha1
sum>
<FOAF:homepage rdf:resource="http://www.jamesbrown.ac.uk" />
<FOAF:img rdf:resource="http://www.jamesbrown.ac.uk/me.jpg" />
<FOAF:research
  rdf:resource="http://www.ontologies.org/acadmicresearch#SocialNetworks" />
<FOAF:research rdf:resource="http://www.ontologies.org/acadmicresearch#Bluetooth" />
<FOAF:knows rdf:nodeID="Jackie Brown" />
 <!--
  More people James know
 -->
<FOAF:weblog rdf:resource="http://www.jamesbrown.ac.uk/blog" />
<FOAF:publications rdf:resource="http://www.jamesbrown.ac.uk/papers" />
 </FOAF:Person>

<FOAF:Group>
 <FOAF:name>Essex University Intelligent Buildings Group</FOAF:name>
 <FOAF:homepage>http://www.essex.ac.uk/intelligentbuildings</FOAF:homepage>
 <FOAF:publications>http://www.essex.ac.uk/intelligentbuildings</FOAF:publications>
 <FOAF:member rdf:nodeID="James Brown" />
  - <!--
  More members of this group
  -->
 <FOAF:research
    rdf:resource="http://www.ontologies.org/acadmicresearch#IntelligentBuildings" />
 <FOAF:research
    rdf:resource="http://www.ontologies.org/acadmicresearch#UbiquitousComputing" />
<FOAF:research rdf:resource="http://www.ontologies.org/acadmicresearch#Agents" />
 </FOAF:Group>
</rdf:RDF>
```