

# SIXTH FRAMEWORK PROGRAMME



Project no. 033841

## EMIL

Emergence in the loop: simulating the two way dynamics of norm innovation

### Deliverable 2.1

## HISTORIES OF THE EXAMPLE OF THE EMERGENCE OF OPEN SOURCE NORMS CASE SELECTION AND HISTORIES

Due date of deliverable: 31.08.07 +45gg  
Actual submission date: 05.10.07

Start date of project: 01.09.06

Duration: 36 months

Lead contractor for this deliverable: UNIS

Project co-funded by the European Commission within the Sixth Framework Programme (2002-2006)		
Dissemination Level		
<b>PU</b>	Public	X
<b>PP</b>	Restricted to other programme participants (including the Commission)	
<b>RE</b>	Restricted to a group specified by the consortium (including the Commission)	
<b>CO</b>	Confidential, only for members of the consortium (including the Commission Services)	

**Contents**

1	Introduction.....	3
1.1	Theoretical strands which have informed the theory of social emergence .....	3
1.1.1	A Brief History of the Concept of Emergence.....	4
1.2	Theories of Social Norms .....	7
1.3	Methodological Considerations.....	9
1.3.1	Activity Theory.....	9
1.3.2	Structuration Theory.....	10
1.3.3	Narrative Analysis .....	13
1.3.4	Speech Act Theory.....	13
1.4	Open Source Literature .....	17
2	Subtask 1: Scope domains for analysis .....	19
2.1	Criteria.....	19
2.2	Case Categories.....	20
3	Subtask 2: Case histories .....	20
3.1	Wikipedia.....	21
3.1.1	Emergence of a new production and distribution paradigm.....	21
3.1.2	How Wikipedia works .....	26
3.2	Analysis of Wikipedia activity.....	38
4	Future steps.....	44
4.1	Extended analysis of Wikipedia Case.....	44
4.2	Wiki Experimental Design.....	45
5	Second Life Case Study .....	45
5.1	Sim Manager Interviews .....	46
5.2	Normative Events .....	47
6	Appendix One: Wikipedia Case Study Code Frame and Definitions.....	48
7	Appendix Two: Wiki Experiment .....	52
8	References .....	53
8.1	Media References .....	53
8.2	Academic References .....	53

## 1 Introduction

The University of Surrey, Centre for Research on Social Simulation is responsible for identifying cases and for the conduct of primary and secondary research into those cases to provide empirical data to inform both theory development and simulator design. The studies conducted as part of this work program are to be informed by and to inform the ontology (EMIL-M) and to provide data as a point of comparison for the simulator (EMIL-S). The intention was also to bring a social science perspective to the study of normative behaviour in the selected cases. More specifically the work package was to:

- Collect data to test the utility of the model from suitable open source scenarios;
- Identify and describe examples of domains where norms for the 'open source' exchange of IPR (e.g. free/libre open source software, academic science, personal photo sharing...) have developed. Thus, explaining particular instances of norm innovation, underlying the main features distinguishing a norm from a simple convention. In all these areas, the institution can only function and prosper if shared norms emerge;
- Use available documentary records, previous research and other secondary sources, for each of these examples, provide a history of the development, examining in particular the emergence and effect of norms;
- Relate these histories to the theoretical approach being developed in WP1 and provide empirical data to inform the models being constructed in WP3.

The work package proceeded by:

- Identifying the major social theoretical strands which have informed the theory of social emergence.
- Reviewing relevant social science concepts of norms.
- Conducting a scan of existing literature on Open Source.
- Identification of alternative methods for the study of behaviour in computer mediated environments.
- Giving consideration to the definitions and concepts in the preliminary EMIL-A.
- Developing criteria for the selection of cases.
- Preparation of recommendation of cases.
- Data collection.
- Conduct of case analysis and writing of case histories.

This report sets out the results of this work to date in this order.

### ***1.1 Theoretical strands which have informed the theory of social emergence***

The concept of emergence has become widely used within the social simulation community. The concept continues to be vaguely defined and to stand in for different propositions about social generative mechanisms. Within the social simulation community, the concept has focused primarily on upward causation (consistent with its usage within complex systems theory and associated research programs such as those into artificial life (Sawyer, 2003)). Few attempts have been made to reconcile this use of the concept with its wider philosophical use and with the parallel debates about the micro-macro link and the relationship between structure and agency within philosophy and the social sciences. Relatively little attempt has been made to identify the defining characteristics of human social systems and to critically re-examine the concept within this context. Similarly derivative concepts such as downward causation and 'immergence' (Castelfranchi, 1998a) have only recently begun to be explored in the simulation of human social systems. EMIL therefore represents a distinctive and rare attempt to advance our understanding of these fundamental social

mechanisms within the context of social simulation. EMIL is concerned to explicate the mechanisms of emergence and immergence within the context of human social systems by focusing on processes of ‘normative’ self-organization. The aim is to contribute both to the conceptualization of emergence as well as to how social emergence may be meaningfully modelled.

### 1.1.1 A Brief History of the Concept of Emergence

The notion of emergence has a long history, having been invoked in a number of disciplines with varying degrees of centrality to the theoretical and methodological development of associated fields. The concept remains ambiguous and contentious, covering:

*...a wide spectrum of ontological commitments. According to some the emergents are no more than patterns, with no causal powers of their own; for others they are substances in their own right... (Clayton, 2006: 14).*

The first explicit use of the concept has been attributed to George Henry Lewes, in 1875 (Ablowitz, 1939). Following Lewes the concept rose to prominence primarily within the philosophy of science but more recently can be seen to have been advanced within three distinct streams: *philosophy*, particularly philosophy of mind; *systems theory*, in particular complex systems; and *social science* where it has largely been referred to under the heading of the micro-macro link and/or the problem of structure and agency. Interestingly there has been relatively little cross influence between these streams.

#### 1.1.1.1 The Contribution from Philosophy of science

The philosophy of science and philosophy of mind stream is arguably the oldest – some date it back to Plato (Peterson, 2006) but the debate is widely seen as having come to focus with the British Emergentists (Eronen, 2004; Shrader, 2005; Stanford Encyclopaedia of Philosophy, 2006). This school sought to deal with the apparent qualitatively distinct properties associated with different phenomena (physical, chemical, biological, mental) in the context of the debate between mechanism and vitalism: the former being committed to Laplacian causal determinism and hence reductionism and the latter invoking ‘non-physical’ elements in order to explain the qualitative difference between organic and in-organic matter. This stream remains focused on explaining different properties of classes of natural phenomena and with the relationship between brains and minds (See Clayton & Davies, 2006 for a recent summary of the positions). Peterson (2006: 695) summarizes the widely agreed characteristics of emergent phenomena within this stream as follows. Emergent entities:

1. Are characterized by higher-order descriptions (i.e. form a *hierarchy*).
2. Obey higher order *laws*.
3. Are characterized by *unpredictable novelty*.
4. Are *composed of* lower level entities, but lower level entities are *insufficient* to fully account for emergent entities (*irreducibility*).
5. May be capable of *top-down causation*.
6. Are characterized by *multiple realization or wild disjunction* (Fodor, 1974) (alternative micro-states may generate the same macro states).

A key concept is *supervenience*: a specification of the ‘loose’ determinisms held to apply between levels such that ‘...an entity cannot change at a higher level without also changing at a lower level’ (Sawyer, 2001: 556). Within this stream prominence of place is given to both downward and upward causation. Clayton and Davies (2006) specify downward causation as involving macro structures placing *constraint* on lower level processes hence ‘*Emergent entities provide the context in which local, bottom up causation takes place and is made possible*’ (Peterson, 2006: 697). This concept appears similar to that of ‘immergence’ within the social simulation literature. This emphasis on the relationship between emergent structure and micro action is an explicit focal point within EMIL and is otherwise largely absent within the approach to emergence typical of complex systems inspired approaches (Sawyer, 2003, 2005).

Davies (2006) argues that the mechanism of downward causation can usefully be considered in terms of boundaries. Novelty, he argues, may have its origin in a system being ‘open’. If novel order emerges it must do so within the constraints of physics. He concludes:

*... top-down talk refers not to vitalistic augmentation of known forces, but rather to the system harnessing existing forces for its own ends. The problem is to understand how this harnessing happens, not at the level of individual intermolecular interactions, but overall – as a coherent project. It appears that once a system is sufficiently complex, then new top down rules of causation emerge (Davies 2006: 48).*

For Davies then, top-down causation is associated with self-organization. For Davies it is the ‘openness’ of some systems that ‘provides room’ for self-organizing process to arise, but he concludes, ‘openness to the environment merely explains why there may be room for top-down causation; it tells us nothing about how that causation works.’ The devil then, is in the detail of the mechanisms specific to particular processes in particular contexts and particular phenomenal domains. Within EMIL the aim is to identify the specific mechanisms that underpin what we call ‘normative’ self-regulation.

### 1.1.1.2 The contribution from Social Science

The micro-macro problem – the relationship between the actions of individuals and resulting social structures and the reciprocal constraint those structures place on individual agency – has long standing in social science. The problem is central to many social theories developed throughout the 19<sup>th</sup> and 20<sup>th</sup> century. Examples include: Marxian dialectical materialism (Engels, 1934) built upon by, among others, Vygotsky (1962) and Lyont’ev (1978); the social constructionism of Berger and Luckmann (1972); Giddens’s structuration theory (1984); and the recent work of critical realists (Archer, 1998; Archer et al., 1998; Bhaskar, 1997, 1998). These alternative theories are frequently founded on differing assumptions, extending from the essentially objectivist/rationalist theory of Coleman (1994), through the critical theories of Habermas and then to the radical constructivism of Luhmann (1990; 1995).

Fuchs & Hofkirchner (2005: 33) classify into four categories the ontological position of alternative approaches to the micro-macro relationship. They argue that the majority of existing social theory falls into one or other of the first two categories: *individualism* and *sociologism*. Neither of these ‘paradigms’ provides a theoretical foundation which supports exploration, let alone an advance in understanding, of the interplay between agency and structure. The third category, *dualism*, was the target of the original emergentists. Only those theories categorized as *dialectical* therefore have relevance. Even here, it is reasonable to conclude that little practical advance has been achieved, as most positions result in a straddling of bottom up and top-down arguments and/or suffer from excessively vague conceptualisation. What has been largely agreed, despite the very different theoretical and often inadequate handling of this problem, is that structure and agency come together in *activity* or in *body-hood* – the specific psycho-motor state at the instant of enaction. Both Vygotsky and Giddens, for example, focus on action as the point of intersection between human agency and social structures and it is arguable that Bourdieu’s *habitus* represents a similar commitment (Lizardo, 2004). This idea has also provided a methodological framework for the study of this type of phenomena, including in computer mediated environments (see below)

### 1.1.1.3 The contribution from Systems Theory

Systems language was clearly evident in the work of the early emergentists and in a great deal of sociology and anthropology – notably that of Margaret Mead and Gregory Bateson. However, ‘systems’ as a focus of systematic research arguably took form with Bertalanffy’s attempt to establish a General Systems Theory (GSM) in 1950 (Bertalanffy, 1950; Bertalanffy\_von, 1968). As the science of ‘wholes’ systems theory stands in contrast to reductionisms concern with parts. In many respects systems theory was put forward as a counter to what was perceived as excessive reductionism dominating scientific discourse during much of the 20<sup>th</sup> century.

While in the early stages of development of the theory, systems tended to be modelled as ‘black boxes’ (effectively masking the relationship between micro and macro elements), the application of the concept to social science, in particular through the development of social cybernetics (Keeney, 1987) and soft systems approaches (Checkland, 1988) provided a theoretical lens as well as methods useful for describing the systemic behaviour of social systems. While the aspiration of GSM to establish a general science of systems is generally regarded to have failed (Jackson, 2000), systems approaches have contributed valuable methods for the study of the interplay between levels. The Systems view of emergence was founded on:

- Holism; the whole is greater than the sum of its parts.
- A concern with *feedback* both positive and negative.
- A concern with boundaries and boundary conditions.

More recently the development of complex systems theory and its application to natural, social and cognitive phenomena has provided additional concepts upon which much current debate about emergence draws. Many of these concepts and methods have become widely used within the multi-agent modelling community (Castelfranchi, 1998b; Gilbert, 1995; Holland, 1998).

Within contemporary debate, and in contrast to the position taken by the British Emergentists who argued that irreducibility was the *exception* (Eronen, 2004), most real world systems are now argued to be non-linear (Kauffman, 2000; Kauffman, 1993, 1996; Stewart, 1990). It is non-linearity which contributes to these system’s capacity for novelty and unpredictability in principle, through the presence of deterministic Chaos (Lorenz, 2001; Williams, 1997) and/or equifinality. Equifinality as it is known within systems theory, or the principle of ‘wild disjunction’ as it is known in philosophy, refers to a system where a single high level property may be realized by more than one set of micro-states which have no lawful relationship between them (Richardson, 2002a, 2002b; Sawyer, 2001). As there is no a-priori basis by which the likely micro state can be determined, such systems are irreducible and unpredictable in principle.

#### 1.1.1.4 Observations

The concept of emergence has led to the establishment of a number of general principles which describe the relationship between micro and macro phenomena, as well as some methods and techniques for identifying and exploring it. Specifically, we can conclude that there are systems which:

- are inherently analytically reducible (to which the concept of emergence does not apply);
- are analytically reducible in principle but difficult to reduce in practice and/or where an advance in science/knowledge is needed for reduction to be possible because the results were ‘unexpected’ (Chalmers, 2006) (to which the concept of ‘weak’ emergence can be applied);
- are not reducible in principle (to which the principle of ‘strong’ emergence is relevant).

There remain some particular challenges both to the conceptual extension of complexity theory and to the use of simulation as a method for the study of human social systems. Natural complex systems demonstrate a capacity to give rise to complex macro pattern solely as a result of *local* interactions between agents. This local interaction can often be characterized as involving some ‘signalling’ between agents. With human social systems, however, this signalling behaviour takes on a qualitatively different form – that of natural language.

The advent of a capacity for language has implications for the behaviour of systems of human agents which we have yet to fully understand. Language makes possible the emergence of domains of interaction which can themselves become the target for further linguistic distinction. This makes possible a range of behavioural plasticity which is unparalleled in natural systems. In addition, or in association with this linguistic capability, humans (and possibly some other primates, cetaceans and elephants) have developed sufficient cognitive capacity to become self-aware – to distinguish ‘self’

from ‘other’<sup>1</sup>. This allows them to conjecture about the effect of a macro pattern on them and others and this may lead them to change their behaviour. This suggests, for example, that an agent can form hypotheses about the relationship between a macro structural aspect of the social system in which it is a participant and then act on that hypothesis, potentially changing the structure which it participates in generating. This gives rise to a feedback path between macro and micro phenomena that is not present (as far as we know) in any other natural system. Understanding these capabilities is fundamental to modelling intra-agent processes (EMIL-A).

A number of level based schemas for emergence have been proposed to deal with this aspect of social systems. Ellis (2006:99-101) identifies 5 orders of social emergence:

1. Bottom up leading to higher level generic properties (examples include the properties of gases, liquids and solids)
2. Bottom up action plus boundary conditions lead to higher level structures (e.g. convection cells, sand piles, cellular automata)
3. bottom up action leading to feedback and control at various levels leading to meaningful top down action - teleonomy (e.g. living cells, multi-cellular organisms with ‘instinctive’ – phylogenetically determined - reactive capability)
4. as per 3 but with the addition of explicit goals related to memory influence by specific events in the individual history (i.e. capable of learning)
5. In addition to 4 some goals are explicitly expressed in language (humans).

This final category equates to what (Goldspink & Kay, 2007) have referred to as reflexive emergence and which Gilbert refers to as second order emergence. It occurs when the agents recognise emergent phenomena, such as societies, clubs, formal organizations, institutions, localities and so on where the fact that you are a member or a non-member, changes the rules of interaction between you and other agents.’(Gilbert, 2002)

The insight behind the EMIL project is that this two-way flow of influence is fundamental to ‘normative action’ in social systems. Agents perceive higher order social structures (norms) and (perhaps) change their (micro) behaviour in response, thus at the same time acting on the norm (perhaps reinforcing it or diminishing it) as a social artefact. EMIL should therefore make a significant contribution to advancing our understanding of this problem: Contributing both to its conceptualization and furnishing methods for its study.

## **1.2 Theories of Social Norms**

The concept of norms has a long history in social theory. Sripada and Stich (2006) open their chapter on a framework for the psychology of norms with a quote from the Encyclopaedia of the Social Sciences: “*No concept is invoked more often by social scientists in the explanations of human behaviour than ‘norm’*”. For Gibbs also ‘*Sociologists use few technical terms more than norms and the notion of norms looms large in their attempt to answer a perennial question: How is social order possible?*’ (Gibbs, 1981). Not surprisingly then the concept has been incorporated into a wide range of alternative and often competing theory arguably finding its peak in the ‘normative sociology’ of Durkheim and Parsons. These alternatives include both psychological (micro) and sociological (macro) theoretical perspectives. All are concerned to explain social (more specifically behavioural) order but alternatively posit norms as a) a structural constraint on human agency or b) as a product of collective agency. Seldom are these two perspectives brought together and even more seldom is the explanatory pathway made explicit.

---

<sup>1</sup> There is considerable debate about whether self awareness depends on language or is a pre-condition for language or both – see for example Ford, P. & Dominey, A. R. n.d. Putting all the Strands Together, Co-evolution of language and theory of mind, [Interdisciplines](#).

The normative literature can be largely divided into two fundamentally distinct groups. On the face of it, these are incompatible; however, in light of the interpenetration of micro-macro perspective discussed above, there may be a basis for their synthesis. In the social philosophical tradition (Lewis, 1969) norms are seen as a particular class of emergent pattern of social behaviour which spontaneously emerges in a population. From this perspective, a 'norm' is identified by an observer ex-post – it is a classificatory distinction, applied when the observer notes some particular characteristic to an emergent social pattern. In the case of a 'norm' that characteristic is the apparently prescriptive/proscriptive character – people behave 'as if' they were following a rule. This has sometimes been presented as an epiphenomenal view of norms. By contrast, the view offered by the philosophy of law sees norms as a *source* of social order. This standpoint assumes the prior existence of (powerful) social institutions and posits them as the source of rules, which, when generally followed, lead to social patterns. A variant on this view further asserts that they need to be explicitly recognised as norms i.e. not just satisfied by unconscious conformity (Castelfranchi, 1998b).

Therborn argues (2002: 868) people follow norms for different reasons. The extremes run from habit or routine to rational knowledge of consequences for the world. Between these lie:

- Identification with the norm or values – linking sense of self (identity) to the norm source (person, organization or doctrine) often leading to in-group-out-group.
- Deep internalization – self-respect – done independently to what others are doing.

Bicchieri (2006: 59) provides a rare hint at the cognitive process involved stating:

*To 'activate' a norm means that the subjects involved recognise that the norm applies: They infer from some situational cues what the appropriate behavior is, what they should expect others to do and what they are expected to do themselves, and act upon those cues.'*

This suggests a complex process of self-classification (how am 'I' situated with respect to this group and what is the nature of the situation in which 'I' find myself, does a norm pertain to 'me' in this situation and under what conditions and to what extent am I obliged to comply?),

There remains considerable confusion as to whether norms are best ascribed as 'in' the environment or 'in' the agent. Within social psychology, the widely adopted theory of Reasoned Action (Fishbein & Ajzen, 1975; Oliver & Bearden, 1985; Sheppard et al., 1988) and the related theory of Planned Behaviour (Ajzen, 1985; Terry et al., 1999) both posit a concept of 'subjective norm'. This is defined as the amount of pressure perceived by an individual to act in a particular way. According to these theories subjective norms act, along with attitudes and efficacy beliefs, to influence agents' intentions and through intentions their behaviour (Bagozzi, 1992; Bagozzi et al., 1989). The extensive empirical research conducted using these theories suggests, however, that this normative pressure has relatively small effect on behaviour (Armitage & Conner, 2001). Smith et al (2007) argue that part of the problem is that these 'subjective norms' are commonly viewed as being 'out there' as '*external pressures which invite compliance*'. They argue for the need to consider individual and social identity as a factor influencing which social norms an agent holds as salient and by which it will therefore be influenced. This is the position of Postmes et al (2005) who argue that the '*processes of self categorisation and identification 'enhance the likelihood that individuals will come to define themselves in terms of a particular social identity...*'. If this is the case, then the norms, stereotypes, and other properties that are commonly ascribed to the group become internalised: they become subjectively interchangeable with personal norms and stereotypes, influencing thought and guiding action. This approach is more consistent with and even suggests mechanisms for the interpenetration of social structure and individual agency which is the focus of our concern.

Nevertheless the specific mechanisms by which norms reflect or reveal emergence/immersion remain ill defined. To begin to identify which (if any) of these loosely defined mechanisms might be supported by evidence and to aid in the development of a theory of norms which helpful for understanding the more general immersion/immersion mechanism, we decided it was necessary

to avoid adoption of any one particular approach and to collect a rich set of data which could be analysed from a range of alternative perspectives to keep our options open. In order to guide the choice some consideration was given to methods advocated within the literature. Approaches identified as particularly relevant to EMIL are discussed below.

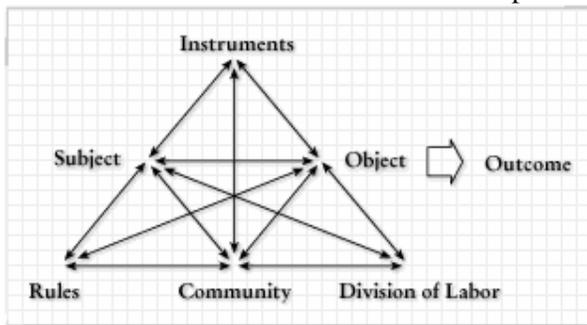
### 1.3 Methodological Considerations

#### 1.3.1 Activity Theory

Engestrom specifically argues for the relevance of Activity Theory in the study of complex non-linear systems particularly in the context of computer mediation.

*To be able to analyse such complex interactions and relationships, a theoretical account of the constitutive elements of the system under investigation is needed. In other words there is a demand for a new unit of analysis. Activity theory has a strong candidate for such a unit of analysis in the concept of object orientated, collective and culturally mediated human activity, or activity system. (Engestrom et al., 1999: 9)*

The following figure shows the essential elements of such a system. It is the internal contradictions or tensions between these elements which provide the basis for the systems dynamics.



In Activity Theory activity is the unit of analysis. An **activity** is a complex process, below it sit: **actions** – specific tasks executed in a time bound manner in order to achieve the object; and routine or automatic **operations**, a particular act in a time and place which does not have an independent goal but serves to adjust an action to a current situation (Kaptelinin & Nardi, 1997). The trajectory of an activity system is described as ‘far from equilibrium’ driven by tensions and contradictions between aspects of the activity system itself. This reflects the Marxian idea that social change is largely the consequence of endogenous conflict or contradictions resulting from the historical conditions and how they play out at a particular time.

In this diagram, *instruments* refer to mediating artefacts. These can be material (e.g. tools, technology) or symbolic (e.g. signs, language, texts, art). The *subject* is the individual or group which is the focus of attention. *Rules*, refers to the norms, procedures and practices relevant to the subject. The community refers to the wider social grouping or ‘community of practice’, to which the subject belongs. *Division of labour* addresses the way in which the activity is carried out by specialist labour while the *object* refers to the product of the activity and furnishes the motivation for the activity. The arrows linking these components of the activity system represent the relationship and it is the quality and nature of this relationship which is the focus for inquiry. Within these lie potentially conflicting forces.

Activity theory recognises two basic processes - internalization and externalization. These operate continuously at every level of human activities.

*Internalization is related to the reproduction of culture; externalization as creation of new artefacts makes possible its transformation (Engestrom et al., 1999: 10).*

*...an individual internalizes, by participating in common activities with other humans, the means of culture: language, theories, technical artefacts as well as norms and modes of acting. (Miettinen, 2006: 392)*

Activity shapes artefacts and this represents an externalisation – whereby the mental processes of the subjects leave a material trace in the form of those artefacts. The particular characteristics of the artefacts subsequently shape the interaction of others and become internalized as mental activity. This is generally cast as tracing a developmental trajectory, whereby ever more abstract objects and concepts are constructed upon those that precede them. Thus societies can construct more and more complex artefacts and individuals more and more abstract means of conceptualisation. These artefacts or tools ‘...connect an individual to other human beings by mediating activity, thereby becoming part of a cultural context’ (Fjeld et al., 2002: 157). Applying this to the collective production of computer software for example, a key artefact is the computer source code. This is an encapsulation of the mental approach to a problem of the producer, alternative solutions capture and allow others to experience and interact as to the relative advantages of alternative solutions, supporting collaboration. The design settled upon will subsequently shape what users can and cannot do, having a material effect on future activity.

It is the focus on the potentially transformative effect of mediating artefacts that has led to a resurgence of interest in Activity Theory, most notably in the computer sciences. Here it is seen as a relevant framework for examining human computer interaction (Fjeld et al., 2002; Kaptelinin & Nardi, 1997; Lewis, 1997) as well as to gain insight into the design of multi-agent systems (Ricci et al., n.d). It is cast not as a predictive theory but rather as a meta-theory which provides a basis for a research methodology. More conventional methods will generally be employed to collect evidence to support any specific analysis.

Engestrom links this idea to Bhaskar as well as to the idea of ‘community of practice’ and Lewis has considered the use of Activity Theory as a tool to explore and analyse distributed learning communities (Lewis, 1997). In this work he highlights the need to focus on selected interrelationships between components of the model.

Kapetlinin and Nardi present a checklist for the application of Activity Theory to human computer interaction:

- a) ‘focus on the structure of the user's activities -- the extent to which the technology facilitates and constrains attaining the user's goals and the impact of the technology on provoking or resolving conflicts between different goals;
- b) focus on the structure of environment --integration of target technology with requirements, tools, resources, and social norms of the environment;
- c) focus on the structure and dynamics of interaction -- internal vs. external components of activity and support of their mutual transformations with the target technology;
- d) focus on development – developmental transformation of the above components as a whole.’ (Kaptelinin & Nardi, 1997: 159)

### 1.3.2 Structuration Theory

As with Marx and subsequently Vigotsky, Giddens was concerned with the relationship between the subjective experience and agency of individuals and social structure (Jones & Karsten, 2003). He proposed that these constituted a necessary duality rather than representing alternative and oppositional paths of explanation for the origins of social phenomena. This can be seen as an attempt to avoid the problems of dualistic approaches identified as unhelpful for dealing with micro-macro interaction (Weik, 2006). As with Vygotsky, Giddens focuses on *actions* as the point of intersection between human agency and social structures. Orlikowski summarises the position as ‘...social actions are situated temporally and contextually, and they always involve interaction between humans. Social structure conditions these social practices by providing the contextual

*rules and resources that allow human actors to make sense of their own acts and those of other people*' (Orlikowski & Robey, 1991: 147). Given this connection it is perhaps not surprising that others have sought to combine Activity and Structuration theory (see for example Widjaja & Balbo, 2005). These authors note however that Activity Theory is philosophically closer to structuralist/functionalist social theory due to its materialist/realist roots, while Giddens is closer to interpretivism. The core concept of *structuration* refers to the process of putting structures into action – it is *'the structuring of social relations across time and space'* (Giddens, 1984). For Giddens though structures exist, not as material entities, but as *'traces in minds'*, being instantiated through action.

Giddens concept of structure is further divided into three dimensions:

- Structures of signification
- Structures of domination and
- Structures of legitimation.

These are associated with corresponding dimensions of individual agency: communication, power and sanction, through the modalities of interpretive schemes, resources and norms respectively. To the extent that actions reproduce existing structures the social system remains stable. As innovation occurs within structures – the system undergoes change or possibly transformation.

Poole discusses possibilities for *'artificially'* bracketing structuration, drawn from Giddens own account – these are consistent with the dual approaches already entrenched within social science i.e.:

- Focus on the individual and take institutions as a backdrop
- Analyse the institutions taking the individual strategic conduct as a backdrop.

In effect what he is arguing for is a combination of functional (external) analysis and an interpretive (internal) analysis – a mixed methodology.

Poole goes on to argue for a methodology embracing the following perspectives:

1. Identify the array of (potential and actual) relevant constitutional structures.
2. Establish relationships between structures.
3. Describe how the social system works by identification of actions that characterize the systems operations and actors interpretive maps.
4. Identify the moves or activities by which agents reproduce structures (Giddens modalities, i.e. interpretive schemes, resources, norms).
5. Identify how structuration influences the context.
6. Account for the roles of factors vis-à-vis one another.
7. Account for the power dynamics underlying structuration.(Poole & DeSanctis, 2002: 15).

Giddens theory of structuration has also given rise to methods developed to examine the role of technology in mediating human activity. However, reading the various interpretations and applications of structuration theory in these papers suggests a problem of ill definition. There are quite clearly a range of alternative interpretations as well as areas of controversy (Poole & DeSanctis, 2002). This observation lends support to prior critiques of Giddens work – that it is vague and difficult to make operational: it provides little methodological guidance (Jones & Karsten, 2003; Poole & DeSanctis, 2002). The variations in use of the theory arise as different researchers use the meta-theoretical perspective in different ways – in particular separate out elements of the theory to support analysis in different ways. To be fair, it is arguably true that all theories which attempt to come to terms with the recursive interpenetration of macro and micro elements have this problem. The moment they are pulled apart their essence is lost and yet, without some level of analytical reduction, they seem to offer little by way of practical support. Such

analysis insists on a temporal deconstruction– the positing of a causal mechanism. As presented by Giddens, the theory provides little guidance as to how this might be done, leaving it open to many alternative approaches. This is however entirely consistent with Giddens own views of social science as ‘irretrievably hermeneutic’.

Unlike Activity Theory, Giddens rejects the idea that structure becomes embedded in artefacts. This presents a problem for its application to the mediating role of technology. This has led to a number of proposed developments of the theory so as to better accommodate the role of technology. Adaptive Structuration theory is one example of this (DeSanctis & Poole, 1994; Jones & Karsten, 2003). Desanctis and Poole (1994: 128) argue that the major sources for structure for groups are the technology, tasks and organizational environment. The influence of technology can be understood by considering its structural features as well as through the values and intent inherent to its design, which they refer to as the ‘spirit’ of the system. The latter is to be interpreted by referring to the signals which reveal the design philosophy including:

- The design metaphor
- Features and how these are named and presented
- The character of the user interface
- Training and help facilities.

Analysis considers the above factors of technology, task and environment, as well as group styles, knowledge and perceptions – examining how they interact. The analysis is pursued at micro, global and institutional level, with discourse analysis constituting a primary method.

The approach proposed by Orlikowski, while on the face of it attends more closely to Giddens concepts, departs from it in spirit (Jones & Karsten, 2003). Orlikowski argues that information technology impacts on several of the modes of structuration. For example, it acts on interpretive schemes as ‘*information technology contributes to the signification order by objectifying and reifying human actors’ knowledge and assumptions, reinforcing them over time*’ (Orlikowski & Robey, 1991: 155). Similarly by making concrete rules and enforcing ways of performing tasks, it may embed systems of domination. It may do this further by imposing certain ways of seeing and thinking – embedding these in its data structures and procedures.

*...information technology will embody the shared meanings, values and goals of [the coalition that built and deployed it] by internalizing and reinforcing the dominant ideology and culture of the organization. In this way information technology can be seen to convey a set of norms that indicate the accepted actions, interests and practices of the workplace. The norms embedded in the information technology constitute moral order, a system of legitimation that directs action and thinking along prescribed paths, and encourages appropriate responses, shared meanings, and common interaction protocols. (Orlikowski & Robey, 1991: 156)*

While framing the argument in terms of Giddens’s theory, this emphasis on material form of structure is more consistent with the objectivist orientation of Activity Theory. This does not necessarily, however, reduce its potential value as an analytical schema.

Both activity theory and structuration theory can account for some level of reflexivity. It is implicit in activity theory that language is an artefact and that abstract artefacts can stand as a basis for the construction of subsequent artefacts. Within structuration theory agents reflexively monitor their activity, although their knowledge may be tacit. Jones cites Giddens ‘*every member of society must know...a great deal about the workings of that society by virtue of his or her participation in it*’ this, he argues leads to a double hermeneutic ‘*whereby the concepts that sociological observers describe are already constituted as meaningful by social actors and can themselves become elements of the actors understanding of their own condition*’ (Jones p 13).

### 1.3.3 Narrative Analysis

Activity theory, structuration theory and their derivatives advance the position that structures are instantiated through activity. The flow of activity provides a means of tracking the historical unfolding – the trajectory of a system. One method consistent with collecting information about sequences of activities, as they are experienced by external observers and through self-observation, is narrative. Narrative is argued to constitute a method particularly appropriate to ‘...*examine the interconnectedness of human agency and social structure and the temporality of historical events in processual ways.*’ (Gotham & Staples, 1996: 481).

Furthermore, narrative construction is argued to be at the core of the functioning of human meaning making – the narrative mode of thought (Bruner, 1991; Dautenhahn, 2002). Bruner observes that ‘...*narrative ‘truth’ is judged by its verisimilitude rather than its verifiability. There seems indeed to be some sense in which narrative, rather than referring to ‘reality’ may in fact create or constitute it...*’ (Bruner, 1991: 13). ‘*Once shared culturally...narrative accruals achieve...‘exteriority’ and the power of constraint.*’ (Bruner, 1991: 19). From this perspective narratives provide an account of how people interpret past events and those interpretations play a role in embedding particular ways of thinking and knowing within the culture – they come to be constitutive of the lived reality. Narrative is a way of connecting the past to the present and provides a basis for future action. When human agents construct narratives they cast themselves as a character, even if it is one of innocent bystander. Consequently, an agents narrative can reveal a lot about the part and future role it may play within a social process.

Narrative data provides insight into the relationship between events – i.e. the emerging dynamics or how events are linked in time. More than this, and significantly for EMIL, as a method it can capture individual and collective accounts of the interplay between individual choices and normative behaviour. It can reveal the agents reflexive account of internal processes (EMIL-A) and of inter agent processes.

### 1.3.4 Speech Act Theory

Any coordinated (social) activity happens primarily in and through language. In many computer mediated environments there may be no other basis for exchanging information and or for influencing the behaviour of others than through computer communicated linguistic exchange. From this perspective Open Source institutions can be viewed as networks of commitments established, maintained and modified, in and through linguistic exchange.

The theory of speech acts views natural language utterances as *actions* (rather than or as well as denotations) on the part of a speaker (Austin, 1962). Speech acts can operate directly or indirectly to achieve ends. These acts may include asking questions, giving orders, or making statements. Austin (1962) argued that all utterances are:

- Locutionary: an act **of** saying something
- Illocutionary: an act **in** saying something; and
- Perlocutionary: having effects upon the feelings, thoughts or actions of the recipients (including the speaker).

In linguistics, this distinction gave rise to the alternative treatment of the semantic (meaning) and pragmatic (consequential) aspects of language. These two perspectives reflect what has become a fundamental concern of language philosophers – the relationship between the word and the world (Eriksson, 1999). It is the illocutionary aspect of speech acts which reveals most about the nature of the relationship between the speaker and the audience. Searle (1969) identified the following classes of illocutionary act :

**Assertives:** commit the speaker to something being the case e.g. suggesting, putting forward, swearing, boasting, concluding.

**Directives:** try to make the addressee perform an action e.g. asking, ordering, requesting, inviting, advising, begging.

**Commissives:** commit the speaker to doing something in the future e.g. promising, planning, vowing, betting, opposing.

**Expressives:** express how the speaker feels about the situation e.g. thanking, apologising, welcoming, deploring, and

**Declarations:** change the state of the world in an immediate way.

Attempts have been made to translate Searle's categories into an axiomatic communication protocol – in particular one suitable as an Agent Communication Language (see for example <http://www.fipa.org/repository/aclspecs.html>). There are, however, philosophical and operational problems with Searle's approach. Searle classifies speech acts based on the pragmatic intent of the speaker. From this perspective '*The speaker (actor) is confronted with an external world which is constituted of objects and states of affairs and is where the listener is also placed*' (Eriksson, 1999). This represents a commitment to a modernist philosophical position (Burrell & Morgan, 1994) positing the existence of a universally available and shared real world underpinning meaning embedded in communicative exchange. Adopting this stance would undermine the possibility of achieving what is sought through this research – an appreciation of the interpenetration of macro and micro phenomena. Modernism has constructed an insurmountable dichotomy between micro and macro where macro phenomena is either regarded as epiphenomenal: resulted from and being reducible to micro phenomena and thus having no (independent) causal power: or reified – treated as concrete and causal primarily with respect to other social structures. Analysis is typically quarantined to one level of analysis in order to avoid entrapment in logical contradictions.

Habermas (1976) proposed an alternative (and intrinsically post-modern) scheme for the classification of speech acts. Whereas Searle assumes the existence of a common shared world, Habermas posits the existence of three worlds: A subjective world of beliefs, desires, intentions and feelings of the actors; a common social world of norms relationships and institutions; and an objective world.

For Habermas, a successful speech act would be one in which the listener both comprehends and accepts the *validity* claims made by the sender and thus enters into the intended relationship. The validity claims include *comprehensibility, truth, sincerity* and *rightness*. Thus for Habermas, a speech act only serves to support the maintenance or emergence of a social institution to the extent that it is held as valid. What is held to be valid is at least in part assessed against other social facts – i.e. conformance with norms or rules – and hence contains a level of self-reference. From this perspective, 'true' meaning is to be found in community consensus not in correspondence to an objective external reality. The 'meaning' of a norm, cannot be found in individual minds but only in the social system to which it relates. Teubner (1989:738) argues that Habermas' approach comes undone as he tries to avoid the implications of this self-referentiality. By comparison, he argues, Luhmann takes self-reference and places it at the centre. He argues that in '*...[Luhmann's] autopoietic systems, discourses cannot but find justification in their own circularity and cannot but produce regularities that regulate themselves and that govern the transformation of their own regularities*' (Teubner, 1989: 737). The Luhmannian position casts modernity as '*...an irreconcilable conflict of different epistemes*'. Self-reference and self-organization then provide a form of stability or closure, while tensions between alternative and conflicting epistemes produce a basis for change (Michailakis, 1995). This is an echo of the basis for change in the dialectic already encountered in Activity Theory. This approach posits society as divided into functionally differentiated 'self-referential' domains maintained in and through communicative acts. The autopoietically reproduced society means that codes specific to a system function only within their respective systems – norms only operate within a domain in which they emerged.

*Society is seen as fragmented into a multiplicity of closed communicative networks. Each communicative network constructs a reality of its own that is, in principle, incompatible*

*with the reality construction of other networks. At the same time, there is a multiplication and fragmentation of individualities that corresponds to the multiplication and fragmentation of social discourses' (Muller, 1994: 52)*

Unfortunately the work of Luhmann (1990; 1995) has a number of problems – not least the displacement of the (thinking, believing, feeling) human actor as the basis for social formation (Mingers, 2002) from the centre of social agency. Both Habermas and Luhmann also fall short of providing an operational framework or at least one with the characteristics necessary for this study. For example, Habermas collapses Searle's speech act categories of 'commissives' and 'declaratives' as well as the normatively authorised 'imperatives' (directives) into the class of 'regulatives' leaving sheer imperatives (directives) as a separate class (Eriksson, 1999). In this way he distinguishes between communicative action and strategic action. The former is action based on consensus (prominent in Open Source) while the latter implies action resulting from the exercise of power or compulsion (largely absent in Open Source). This is helpful from the perspective of the study of norms and rules, however, while solving the problem of unhelpful conceptual entanglement it results in a problem of insufficient resolution in the critical area of the category of 'directives' fundamental to normative action.

The intrinsic openness of Open Source and the consensual nature of associated communities, in particular the common absence of formal means of compulsion or sanction mean that the majority of exchanges will be communicative acts – i.e. bounded and influenced by normative behaviour rather than through the exercise of formal authority, power or coercion. The existence of 'community' may be central to establishing such an environment as the diversity of social backgrounds and experiences participants bring to a task may fail to have sufficient power to provide coherence. It is important, therefore, to be able to discriminate between these subtle influence behaviours and Habermas' taxonomy is weak in this area. Eriksson (1999) has proposed a synthesis of both Searle's and Habermas' positions which adds categories back in this critical area while maintaining the mapping between them. However, neither Searle, Habermas, Luhmann nor Eriksson provide a classification scheme of natural speech acts which is capable of capturing the many subtle ways in which people attempt to influence one another through their utterances – nor what this subtlety may reveal about the depth and quality of their relationship.

A variety of social psychological theories have been developed for measuring the quality of relationships in communication. Meissner (2005) groups these into three categories of which only the third (relational topoi) are relevant to the needs of EMIL. In contrast to Speech Act Theory, systems designed to measure these relational topoi have been derived either empirically (Walther, 1995; Walther & Burgoon, 1992) or through a combination of principles and empirical observation (Stiles, 1992). Many of these schemes are domain or context specific and unsuited to classification of a wide range of speech acts.

The Verbal Response Mode taxonomy (Stiles, 1992) has its origins with research into the relationship between psychologist and client. It is designed to operate with dyadic one to many or many to many relationships. Despite its narrow origin it has been developed over many years and used in a wide range of communication contexts. Stiles defines it as '*a conceptually based, general purpose system for coding speech acts. The taxonomic categories are mutually exclusive and they are exhaustive in the sense that every conceivable utterance can be classified.*' (Stiles, 1992: 15). The classification schema has attributes very attractive where there is a need (as in the EMIL project) to capture many of the subtleties of natural language use that derive from and rely on the intrinsic flexibility (and ambiguity or soft edged nature of natural language classification) to a more formal or axiomatic system needed for simulation.

VRM uses three structural (rather than functional) principles. The three principles are a concern with whether the speaker frames an utterance on the basis of:

1. his/her own or the others source of experience

2. presumptions the speaker makes about the others experience (feeling, perceiving or intending); and
3. Whether the speaker presents the experience from his/her own viewpoint or a viewpoint shared or held in common with another.

These *‘three principles are dichotomous – each can take the value ‘speaker’ or ‘other’ -and they are orthogonal in the sense that all eight (2\*2\*2) are possible’*(Stiles, 1992: 15).

It can readily be seen how this scheme is consistent with the three worlds of Habermas. Principle one involves a choice of example drawn from own or others world. Principle two involves the making of presumptions about the world of the other and principle three, ‘choice of frame of reference’ involves reference to ones own world or a shared social world (e.g. to a social construct such as a norm or a shared objective world where it relates to a concrete entity). The shared or co-constructed world acts as an important bridge, as Stiles notes *‘Strictly speaking, in order to understand their own utterances, speakers cannot use a frame of reference that is exclusively the other’s. Consequently the taxonomy distinguishes between use of the speaker’s personal frame of reference and use of a frame of reference that is shared with the other’* (Stiles, 1992: 62)

Many of the ways in which we use speech to influence others (its illocutionary force) is conveyed through subtleties in the way something is phrased rather than overtly through what is said. Each of the VRM categories captures the micro-relationship by which speakers are linked in an individual utterance. The coding of the illocutionary force is performed at two levels – that of literal meaning (form) and pragmatic meaning (intent). Coding the literal meaning implies attending to what the utterance says *‘based on the dictionary meaning of the words and the standard meaning of the grammatical construction’* (Stiles, 1992: 65). The pragmatic meaning is what the context of the utterance and its literal meaning reveals about the speakers intent – what he or she intends as an outcome of the exchange in terms of the behaviour of the other or the effect on the relationship.

Coding both literal and pragmatic meaning against each of the eight dimensions yields 64 possible combinations or modes. While this is better than many coding schemes it cannot capture some of the more subtle aspects of human communication - in particular what Stiles refers to as ‘off-record’ meanings such as ‘hints’ at an action or manipulations. Nor does VRM code the denotative or connotative meaning so to the extent that what is being talked about is important it requires separate coding.

Using the above principles all utterances can be assigned a unique code. This code is classificatory and does not require a judgement to be made about degree. However a gauge of ‘force’ is available through the resulting modes. The modes are defined by the points of intersection of the three dimensional matrix resulting from application of the three principles. The VRM terms used to label the modes (terms such as disclosure - see below), ‘overlap the centre of the natural language’ meaning of that term well and this is helpful in coding – avoiding the need for the coder to learn a new and unfamiliar technical definition.

The VRM has been used in research into a wide range of communication contexts. From it have been derived a range of indexes about, for example authority, politeness and attentiveness (Stiles, 1992; Stiles et al., 1997). One clear finding is that the distribution of modes used in an exchange differ markedly between different roles, different relationships and among verbal tasks. A relationship characterised by a power difference (authoritative /deferential) is quantitatively identifiable through the relative incidence of alternative modes.

The VRM thus has many attributes that make it attractive as a basis for analysing communicative acts associated with EMIL. These include the fact that it provides a clear and comprehensive categorisation scheme which, while able to capture some of the subtle ways in which language is used to influence, uses well bounded dichotomous/orthogonal categories rather than overlapping interval variables. This simplifies coding and lends itself to formalisation and thus incorporation into an Agent Communication Language as part of a computer simulation. This potential for

formalisation has been demonstrated by recent attempts to automate coding of transcripts using VRM (Lampert et al., 2005). These have led to reasonable levels of reliability.

### 1.4 Open Source Literature

The World Wide Web has made possible new production processes that are global in scale. A new production model has resulted built around an ‘architecture of participation’ (O’Reilly, 2004) or a so called Bazaar Governance (Demil & Lecocq, 2003; Raymond, 2001). By exploiting the very low transaction costs of web media, the possibility exists to attract a small contribution from a very large and diverse group of contributors to develop information intensive products and services. This is the general model which is argued to have made ‘open source’ possible. The Open Source movement is attracting attention as it represents a form of self-organization of social and productive activity quite different to both command and control and market based governance systems (Demil & Lecocq, 2003; Raymond, 2001). Open Source arguably lacks or has reduced reliance on the more common social ordering mechanisms, relying instead on self-organization across distributed networks (Muffatto & Faldani, 2003). The evidence makes clear that the mechanisms upon which open source relies were not pre-conceived but rather have been discovered through practice and are still ‘in process’ (Raymond, 1998). Theory is struggling to catch up with and to explain the phenomena (see Rossi, 2004 for an overview). Understanding it is important if we are to appreciate how small local contributions can be effectively brought to bear on large scale global problems. How can the loosely orchestrated contribution of hundreds or even thousands of actors self-organise in order to generate a solution to some target issue or problem?

The scan of the literature revealed that there are alternative and conflicting views about the fundamental ordering mechanisms which support Open Source collaboration. In one of the more comprehensive reviews of the literature, Rossi notes ‘*Scholarly contributions ...tend to reflect author’s personal conception of human nature and their beliefs about the interpretive power of particular theories of social interaction...*’ (Rossi, 2004). In particular theorists are divided about whether the order is to be primarily explained by intrinsic motivation or extrinsic motivation.

Intrinsic motivation is variously attributed to altruism (Hars & Ou, 2002) group membership (Bowles & Gintis, 2003) social identity (Dholakia et al., 2004; Ren et al., 2007) and associated feelings of obligation and conformance and associated reciprocity (Fehr et al., 2002), gifting (Bergquist & Ljungberg, 2001; Zeitlyn, 2003) and community membership (Edwards, 2001). Other theories include political/ideological beliefs (Escher, 2004), either positive (that software should be free) or negative (that Microsoft should be opposed), commitment to ‘freedom of choice’ (Elliott & Scacchi, May, 2004) or to the values of the ‘hacker culture’ (Chance, August, 2005; Lin, 2004). Yet others suggest enjoyment, learning and challenge (Luthiger Stoll, 2005) as key motivators including the psychological experience of flow (Novack & Hoffman, 1997). Many of these alternative positions incorporate a concept of normative action and all have implications for both inter and intra agent mechanisms and processes and hence for EMIL.

Theories of extrinsic motivation draw primarily on theories of the economic rational actor or transaction cost economics. Here voluntary contribution to Open Source is attributed to rational utility maximisation, with the utility being attributed to a range of short and long term benefits, including skills and/or ‘reputation’ (Watson, 2005) which may be traded for employment opportunity. Other variants include peer recognition and prestige (Johnson, 2002) or the ‘economy of regard’ (Dalle et al., 2004). Open source is variously cast as a response to heterogeneous need (Franke & von Hippel, 2003) or public good (Christley et al., 2004). Again, with the exception of the crudest of utility maximisation theories, these approaches often embrace some concept of norm and have some relevance to EMIL.

As Rossi notes however, ‘*...the question whether one or other category of motivation dominates is not as interesting as the question concerning how the different motives interact.*’. There is work which explores such interaction, Osterloh et al (November, 2002) for example are concerned with the interplay of motivational, situational, and institutional factors. They argue that the new

production model characteristic of OSS depends on a well balanced portfolio of intrinsic and extrinsic motivation, low costs for contributors and governance mechanisms that do not crowd out intrinsic motivation. This approach supports analysis of OSS using institutional theories or theories of community (Pedersen, June, 2002; Preece et al., 2003). Despite these efforts however, the process is poorly understood and the potential exists for EMIL to make an important contribution in this area.

## 2 Subtask 1: Scope domains for analysis

Cases needed to be identified which would support analysis at both micro and macro levels. Case selection was undertaken following four steps involving all partners to the project.

- Establishing criteria for assessment.
- Identifying broad categories of cases.
- Canvassing examples of categories.
- Evaluating and selecting examples against the criteria.

### 2.1 Criteria

Criteria identified as important for the selection of cases included.

- Timing/availability (research able to be conducted over a time frame consistent with the project plan)
- Accessibility of data
  - By-product data (email lists, bug reports, logs etc.)
  - Directly by electronic means or one-on-one.
  - Cases with freer access (geographically and otherwise – i.e. no good if key players are institutionally quarantined due to political sensitivities or similar) to be preferred
  - Non reliance on ex post data
  - Open to application of methods sensitive to the required social (normative) data.
  - Open to application of methods sensitive to the required data at both macro and micro level)
- Complexity (Moderate to high task complexity more likely to generate ‘norms’ than simple protocols; but more difficult to model. Prefer cases which represent continuum of complexity but requiring similar model morphology)
- Size (Too many different agent *types* and the model becomes difficult to manage intersecting networks (internal clusters of agent dynamics) increases model complexity. Complexity of the task increases with interplay between formal and emergent aspects of institutions. Complexity of the task increases in presence of some form of technology mediation)
- Self-organization evident
- Interesting to wider community
- Clear example of norm innovation going against the current state of affairs.

As the selection of cases was undertaken concurrently with the development of the EMIL-M as per WP1, explicit and close linkage between the two was not pursued at this time. It was also seen as desirable to pursue rich cases and to adopt an exploratory methodology which avoided making too many assumptions about the nature and form of normative action within open source communities.

## 2.2 Case Categories

Four categories of cases were then identified (including that of Open Source which was regarded as mandatory) and evaluated against these criteria.

Case type	Advantages	Disadvantages
Cyber 'worlds' (eg Secondlife)	Already 'simulations' with self selection of salient variables and simple environments. Relatively quarantined from real world due to actors operating 'in role' – reduced need for 'real' world contextual information about participants. Self-organising social systems.	Unknown quantity in terms of access to members. Could be time consuming to study (ethnographic immersion). Access to some aspects of micro data (motivations and intent) may be difficult to infer or obtain by direct means. Relevance to real world behaviours would need to be established or argued.
Open source projects	Target as set out in funding bid. Attracting international interest as an example of self-organised sociality. Clear markers of self-organization (power law).	Task environment may be somewhat 'tame' due to modularity and analytical tractability of software development. Social complexity may be low (due to granularity - market rather than society?) Main projects of interest (large/complex) difficult to get micro data on – resistant respondents) Need real contextual data to inform model (existing research makes clear that agents influenced by real context as well as online context)
On-line mediated communities of practice (eg Nemo)	Deal with wicked problems By-product data readily available to test fit at macro level. As a group of paid professionals may be open to participating in research and responding to requests for data at micro level	Need for supplementary data to calibrate micro aspects of model (motives and intent, other real world factors – such as interests outside online context) Not all players accessible to equal degree (lurkers).
Real world communities of practice.	Deal with wicked problems. Data should be accessible using standard methods at both macro and micro levels. As group of paid professionals may be open to participating in research and responding to requests for data at micro level (unless already 'over-studied' e.g. BP)	May be structured on semi-formal lines – not fully self-organising. Operate in very complex environments increasing the level of model detail and presenting calibration challenges.

## 3 Subtask 2: Case histories

Examples of these categories were canvassed from all partner organisations and further considered against the criteria. Cases selected following discussion were:

- Open Source Software Development
- Second Life.

The original intention was to focus on Open Source Software Development. A literature review was undertaken and some initial investigation was made into the nature and type of transactions evident within the developer databases in particular evidence of 'normative' interaction. This investigation revealed that behaviour within most of these developer discussions was stable and that evidence of normative interaction (particularly norm innovation) was rare and diffused. Consideration was therefore given to alternative 'Open Source' projects.

### 3.1 Wikipedia

Wikipedia is attractive as data is readily available in the form of archives, history lists and discussion pages as well as the article pages. A preliminary scan revealed evidence of some evolution of rules and etiquettes as well as communicative exchanges which both adhered to and violated those etiquettes. It was judged that the volunteers that have participated in creating the Wikipedia have emerged a set of permissions, obligations, rules and norms which bring it into being and maintain it as an institution. From a governance perspective there are relatively few means within Wikipedia by which formal control can be exercised and the community must rely instead on the use of informal or 'soft' control. These mechanisms need to be effective in the face of perturbation from 'vandals' (task saboteurs), 'trolls' (social saboteurs), as well as turnover of contributors in the context of a task which requires the accommodation of emotionally charged and value based issues.

From its beginning in 2001 when the only rule was 'there are no rules' (Sanger, 2005), through to 2007, Wikipedia has amassed a substantial array of rules and guidelines. Most have been made explicit as the community struggled to deal with its exponential growth (Viegas et al., 2007). The need for, the nature of, and mechanisms for enforcement of these rules has, however, been a very controversial aspect of Wikipedia's development (Sanger, 2005, 2007b). Wikipedia can, therefore, assist us better to understand:

- Changes in the use of norms and rules in self-organising communities as task complexity increases over time;
- The range and type of rules and norms used to self-regulate open volunteer communities where there is little to no hierarchy and limited capacity for formal sanction;
- How these norms and rules are invoked and maintained through communicative and administrative acts;
- The relationship between goal, technical artefacts and social structures and the exercise of individual agency.

A comprehensive review of literature pertaining to the Wikipedia was undertaken and the findings are summarised below.

#### 3.1.1 Emergence of a new production and distribution paradigm.

Encyclopaedia development is enormously time consuming with most internationally recognised encyclopaedias representing sunk investment developed over decades and needing only revision to maintain. From the perspective of conventional business strategy there were, therefore, very substantial barriers to new entrants. In their book 'Blown to Bits', Evans & Wurster (2000) use the Encyclopaedia Britannica as an example of the impact of new media can have on such barriers.

Begun in 1768, the encyclopaedia Britannica became a benchmark; regarded as the most 'comprehensive and authoritative' of all general subject encyclopaedias. Since its inception it underwent 15 editions (revision being undertaken every 4 or 5 years). It was an expensive and exclusive product and Evans & Wurster note that accordingly, direct sale was used for most of its history, with sales peaking in 1990. At the time of their analysis (2000) those sales had collapsed by 80%, being displaced by CD-Rom based competition selling for around 3% of the price. New media effectively cut the cost of distribution and information was on its way to becoming a mass product.

Britannica responded late to this first new media threat and then found that what had once made them most competitive – their comprehensive coverage – had become a threat when CDROM media proved restrictive given the large content. Moreover, the existence of a large direct sales force was inconsistent with the distribution requirements for a CD-ROM. In a classical straddling compromise (Porter, 1996) the text only CD was bundled with the paper version and still distributed by the traditional means. A CD only version of Encyclopaedia Britannica still sold at around \$1000; 200%

above the competition and, being only text based, failed to exploit the potential of the media. The company struggled to retain its market position and was sold in 1996 for half of its book value.

At that time the competition, such as that from Microsoft with its Encarta product, often became established by buying basic content cheaply from less competitive paper based encyclopaedia developers and adding archival (non-copyright) multimedia content to better exploit the capabilities of the new media. Thus they avoided having to begin from scratch. During this period CD-ROM was of course itself displaced by the Web. In an attempt not to be caught out again, Britannica was the first encyclopaedia to place its entire content on the web (Disabatino, 2001), commencing in 1994, with access on a subscription basis. However, the company changed its strategy 3 times between 1994 and 2001, moving from subscription to free and then returning to subscription as it struggled to find a business model that would work in the new environment. Following the collapse of the dotcom bubble, and at the time it reverted to a fee-based model, Britannica argued that the days of free content were over. This position was directly contradicted at the time by then Nupedia/Wikipedia founding editor Larry Sanger (2001). Sanger argued that far from being over *'The grandest days of free content have not yet begun'* (Sanger, 2001). So far the evidence is in support of Sanger.

Wikipedia represents much more than an alternative source of general information – it is an example of a radical new production model built around an 'architecture of participation' (O'Reilly, 2004). By exploiting the very low transaction costs of web media, the possibility exists to attract a small contribution from a very large and diverse group of contributors to avoid the entry problems that used to confront the development of information intensive products and services. This is the general model which, in retrospect, is argued to have made 'open source' possible. As a production process however, the evidence makes clear that it was not pre-conceived or invented but rather has been discovered through experiments which often had their rational beginnings in traditional models. Wikipedia is very much an example of this.

### 3.1.1.1 Origins of Wikipedia.

Wikipedia grew out of Nupedia. Nupedia was initiated and founded by Jimmy Wales, underwritten by the Dotcom start-up Bomis.com. Larry Sanger was appointed as its first editor-in-chief. From its inception Nupedia was linked to a free information concept. It switched from a home-grown license to the GNU Free Documentation License in January 2001 and thus joined the wider open source movement. It was not alone: Richard Stallman started the GNUPedia project at much the same time. Both Nupedia and GNUPedia adopted a free content license but used traditional 'expert' methods from compiling content. The process Nupedia used to review submissions comprised seven-steps:

1. Assignment
2. Finding a lead reviewer
3. Lead review
4. Open review
5. Lead copy-editing
6. Open copy-editing
7. Final approval and mark-up (Nupedia, 2000)

Nupedia contributors were expected to be experts; the policy stating, *"We wish editors to be true experts in their fields and (with few exceptions) possess [Ph.D.s.](#)"* Wider qualifications were accepted, but checks on claims were conducted and most contributors were academics. The complex and time consuming review process and lack of openness have been argued to be the reason for the failure of the Nupedia project. Sanger has, however, questioned this view (Sanger, 2006b, 2007a), arguing that the expert model was sound but needed to be simplified. Significantly,

at least in the very early stages, the technology used to build, discuss and approve content was cumbersome.

Sanger was introduced to the WikiWiki software platform in 2001 and saw in it a way to address the technological limitations hampering Nupedia. The inherent openness of the Wikiwiki environment was, however, seen as a problem and as a consequence Wikipedia began as an experimental side project to Nupedia. The initial idea was to allow collaboration on articles on Wikipedia prior to, or as a part of, the peer review process for inclusion in Nupedia. The intrinsic openness of Wikipedia attracted increasing numbers of contributors and it quickly developed a life of its own, functioning independently to Nupedia. At this time the dotcom bubble burst, Bomis.com got into difficulty and Jimmy Wales discontinued funding for a salaried editor-in-chief in December 2001. Sanger remained in a voluntary capacity for a short period before resigning from both projects. The Nupedia website was shut down on September 26, 2003. In 2006 Sanger stated *'Nupedia was allowed to wither untended. It could have survived; its main logistical problem was that its editorial process simply had too many steps. Bomis essentially cut the funding to Nupedia just before we were adopting a new, much more efficient system.'* (Sanger, 2006b). Sanger notes also that a majority of the Nupedia Advisory Board did not support the Wikipedia: they being committed to 'rigor and reliability'. He states that some of them *'... thought that a wiki could not resemble an encyclopaedia at all, that it would be too informal and unstructured'* (Sanger, 2007a).

At the time of his departure Sanger was still predicting a bold future for the free encyclopaedia. At that time Nupedia had just over 20 entries which had been compiled over 18 months (Sanger, 2007a). By contrast, the open access Wikipedia, supported then by just a few dozen regulars and a growing collection of add-hoc contributors, had amassed 6,000 articles in the first six months (Sanger, 2001).

From Sanger's earlier comments it is clear that he had been surprised at the rate of development and of the quality achieved on Wikipedia by the relatively un-coordinated action of many editors. His vision at that time was, and remains, for an open environment feeding into an expert based quality assurance process. Just before he left he said:

*...to paraphrase Linus Torvalds, "Given enough eyeballs, all typos, factual errors, and other errors of content are shallow." ... Wikipedia articles can be vetted further by Nupedia. As explained above, Nupedia's principals are, even now, setting up a system whereby the best content from Wikipedia can undergo Nupedia's rigorous review process. Time will tell, but the combination of Nupedia and Wikipedia, together with the new Chalkboard, seems to be an unstoppable high-quality article-creation juggernaut. Many articles will begin life on Wikipedia as mediocre, short items, where they will be expanded and cleaned up; the best of these will move to the Chalkboard, where Nupedia's experts will further hammer them into shape; then, by the time they enter Nupedia's system, there will be almost no work to be done on them, and they will slip through the system very quickly. This too is a prediction, but it seems reasonable enough. (Sanger, 2001)*

At that time, Sanger noted that it was conceivable that Wikipedia could continue to produce articles at a rate of 1,000 per month. He noted also that at that rate, in seven years, there would be 84,000 articles (Sanger, 2001). He foresaw, however, that growth beyond this was possible as the project attracted increased attention and snowballed. While his vision of a Wikipedia feeding Nupedia was not realised, his prediction for growth most certainly was. As at April 2007 (i.e. less than 6 years later), there were over 1,734,076 entries in English and there were 75,000 active contributors working on 5,300,000 articles in more than 100 languages (source: Wikipedia). Growth by number of edits and numbers of articles has been exponential (Voss, 2005). In their paper on the evolution of Wikipedia, (Almeida et al., 2007) note that growth in articles, editors and users have all shown an exponential trajectory. Growth in articles and edits has been due primarily to the growth in the number of contributors contributors. Overall productivity of contributors has tended to fall over time.

Approximately 70% of Wikipedia contributors do not create any article but only contributed to existing articles. *'The burden of creating articles is concentrated on only 30% o contributors'*. This has remained largely constant over its history. Level of contribution varies markedly. Forty percent of contributors have made updates to at least 4 articles while 1% has contributed to more than 1000. (Almeida et al., 2007)

### 3.1.1.1.1 *The Debate over Governance*

Wikipedia is continuing to attract a great deal of interest. This appears to be due to it often being judged (despite its ongoing critics) to have achieved a level of quality well beyond what might be expected given its open character. Sanger remains one of its critics, still arguing for a hybrid model. He has, for example, argued that the success of OSS projects was achieved by balancing the need for open involvement and technical excellence (Moody, 2006; Sanger, 2004, 2006a, 2006c). Recently he stated *'Wikipedia has shown fantastic potential, and it is open content--and so if the project has problems (or features) which will keep it from being the maximally authoritative, broad, and deep reference that I believe could exist, I firmly believe that the world has the right to, and should, improve upon it.'*(Sanger, 2007a).

This is not a view shared by the other co-founder Wales (Wales disputes that Sanger was a co-founder, see Mitchell, 2005) who's original vision was for the project to be completely open (Sanger, 2007a).

Some of the key principles which still operate for Wikipedia were initially developed at the foundational stage of Nupedia. Sanger says for example... *'one of the first policies that Jimmy and I agreed upon was a "non bias" or neutrality policy ... rejecting an alternative that ... some early Nupedians plugged for: the development, for each encyclopaedia topic, of a series of different articles, each written from a different point of view'*(Sanger, 2007a).

The use and enforcement of rules has, however, been an ongoing issue within the Wikipedia community with an apparent division emerging both between the founders and within the wider community about whether rules were necessary and if they were, how extensive they should be and how they should be pursued. Records of this debate are to be found on the [WikiEN-L](#) mailing list.

Initially much of this debate was undertaken on the Wiki itself. A number of attempts were made to deal with this by modifying the characteristics of the environment and to turn what was becoming a problem into a potential asset. One of these was to create the meta-wiki – a forum for discussion of matters pertaining to the Wikipedia environment, rules and administration, rather than issues of content. Sanger notes, however, that this became even more uncontrolled. Sanger and, following his departure Wales, were the only administrators with the power to exclude participants from the site. In 2004 this authority was passed to an Arbitration Committee and to people granted administrator status. Within the English Wikipedia (different systems have evolved in different language sites) prospective administrators can be nominated by users (including self) and, provided they receive support from around 80% of voting members (Riehle, 2006), may be granted administrative privileges. The Arbitration Committee is seen as a mechanism of last resort in the dispute resolution process and only deals with the most serious disputes and cases of rule-breaking. It alone can impose binding solutions up to and including a ban from editing Wikipedia (source Wikipedia). Recommendations for appointment to this committee are made by open elections with appointment the prerogative of Wales.

In 2002 Sanger states *'It seemed that participation in the community was becoming increasingly a struggle over principles, rather than a shared effort toward shared goals.'*(Sanger, 2005). There was active resistance by part of the community to any attempt to enforce rules and, according to Sanger, community sanction was proving insufficient to control inappropriate behaviour. This led to the spirit of collaboration being displaced by a spirit of competition. *'It is one thing to lack any equivalent to "police" and "courts" that can quickly and effectively eliminate abuse; such*

*enforcement systems were rarely entertained in Wikipedia's early years, because according to the wiki ideal, users can effectively police each other. It is another thing altogether to lack a community ethos that is unified in its commitment to its basic ideals, so that the community's champions could claim a moral high ground.'* (Sanger, 2005).

In November 2002 Wales, despite his avowed commitment to openness, placed some principles outside debate.

```
O.k., I hereby proclaim the following:
> * We will not tolerate biased content. The neutral point of view is not
> open to vote; it's decided. If you don't like it, go somewhere else.
>
> * There are certain other policies as well that basically define us as a
> community. We have arrived at them by broad consensus, and they should be
> respected. Wikipedians working in good faith should feel empowered to
> enforce those policies. They shouldn't have to apologize for doing so!
>
> * We will not stop banning vandals. We should seek out the best ways we
> know how to make sure that non-vandals are not lumped in with the vandals,
> but please stop talking as if we'll just stop banning them, because it
> ain't gonna happen.
>
> * We try to help newcomers who want to contribute but don't quite
> understand the body of good habits (and rules) we've built up. But we
> should not and *will* not tolerate forever people who are essentially
> attempting to undermine the system. See below.
>
> * To whatever extent we are or are not, or should be, a democracy, the
> following is also true. We are a benevolent monarchy ruled by a
> "constitution" or, anyway, a developing body of common law that is not
> open to interpretation, but not vote. This has been the case from the
> beginning, and we aren't going to change that. Source (
http://lists.wikimedia.org/pipermail/wikien-1/2002-November/000086.html)
```

Sanger argues that the governance problem emanated, at least in part, from their failure to realise that the Wikipedia would form a community rather than being 'just an encyclopaedia project': He states '*...for a community attempting to achieve something, to be serious, effective, and fair, a charter seems necessary.*' (Sanger, 2005). His position is that, had they appreciated that Wikipedia would become a community; they would have designed a set of governance mechanisms to frame the coordination process while still allowing rules of content management to emerge from the community so created. What were missing were the mechanisms for dealing with the inevitable disputes about governance principles themselves. The evolution of rules can be traced through the archives – see <http://en.wikipedia.org/wiki/Wikipedia:Archive/RulesToConsider>

Sanger remains of the view that the governance problem within Wikipedia is serious and debilitating and that it compromises the integrity of the project. Sanger has now established a new project, Citizendium, which incorporate principles and rules of governance as well as incorporating a process of expert review in accordance with his original vision (Sanger, 2007c).

The argument between those who advocate the need for tighter formal control and a capacity for sanction for a project such as an encyclopaedia and those who advocate a more libertarian orientation consistent with the values of the 'hacker' culture and 'wikiwiki way' is ongoing and still divides contributors and commentators. This debate is, however, largely contained to online and academic discourse.

Within the public domain the debate is more about the quality and reliability of the content of Wikipedia entries. There is some concern also about libellous entries and bias. There is some evidence that this concern may be being fed, if not orchestrated by Wikipedia competitors including Encyclopaedia Britannica – reflecting a commercial rather than a philosophical difference.

### 3.1.2 How Wikipedia works

#### 3.1.2.1 Wiki technology

The defining characteristics of Wikipedia need to be understood as a starting point to the analysis of it as a social phenomenon. Lih (2004) describes the wiki as ‘social software’ having a feature set which actively supports communication and collaboration. For Holloway et al (n.d) also, it ‘...*relies on facilitating human interaction rather than superseding it*’, arguing that ‘*Encyclopaedic content is so complex that a process for reasoned discourse is the only practical way to reach agreement.*’

It can be argued that the main impact of this technology is the way in which it alters transaction costs (Coase, 1993; Williamson & Winter, 1993). Hodgson (1996) has argued that transaction costs are essentially costs associated with managing information. Increasing transaction costs result from increased division of labour and technology mediation can help mitigate these costs, increasing the number of individuals who can effectively collaborate to solve problems. According to Ciffiolilli (2007) ‘*Wiki technology in a way literally cancels transaction costs for editing and changing information.*’. However, the technology does not cancel other costs of coordination.

Costs associated with the need to focus activity and ensure that it is relevant to the task (producing a quality encyclopaedia for example) can be considered as governance related costs. The highly open nature of the wiki – the very characteristic that reduces transaction costs associated with editing and information updating – may increase costs of governance. There may well be design tradeoffs to be managed. As technology and culture combine to influence transaction related costs, this trade-off involves how best to mix the use of technology imposed control with cultural control. The efficacy of cultural control will itself be influenced by factors such as the homogeneity of the user group and that group’s propensity for self-organisation (endogenous norm formation), rates of turnover of the group, and the effect of external perturbation of the group or of the task on which they are working. Heterogeneity of goals, motives, values and styles may all be relevant to ascertaining what balance of technological and cultural artefacts may combine to best match available talent to the task. What then are the defining characteristics of the wiki artefact?

Wiki technology has a **very flat learning curve**: contributing is extremely simple. There are few technical impediments (and hence low transaction costs as well as costs of contribution) confronting novice users.

Wiki platforms are intrinsically open and support decentralised action unless modified to control or restrict access. **Division of labour emerges** as editors choose which pages interest them and which they want to focus on contributing to or maintaining.

Wikipedia has added a number of facilities which support the ready detection and correction of vandalism of pages. These include watch lists, history lists and version comparison and reversion facilities. Watch lists also support users in taking responsibility for the oversight and monitoring of particular topics. **Watch lists** provide a mechanism whereby an editor may request an email when a page they are monitoring is changed.

Changes made to a page are logged using a **history list** which supports comparisons to be made between versions as well as identifying the time and date of any change and the ID of who made that change (IP address in the case of a non registered editor). The reversion facility contributes a number of distinctive benefits. The most obvious of these (and that for which it was no doubt designed) is to support the rapid reinstatement of the page content prior to an act of vandalism. Lih (2004: 4) attributes significance to this feature noting that ‘*This crucial asymmetry tips the balance in favour of productive and cooperative members of the wiki community, allowing quality content to emerge*’. It may have benefits beyond this, however, as Stvilia et al (2004: 13) note that ‘*By allowing the disputing sides to obliterate each others contributions easily, a wiki makes the sides*

*interdependent in achieving their goals and perhaps surprisingly may encourage more consensus building rather than confrontation’.*

Stvilia, Twidale, Gasser, & Smith (2005) among others identify **discussion pages** as an important ‘*coordination artefact which helps to negotiate and align members perspectives on the content and quality of the article.*’ Discussion pages provide an opportunity for managing minor disputes about content or editing behaviour and for movement towards the agreement or consensus needed to for the production of an encyclopaedia. This need is distinctive to on-line collaborations which have as their intended outcome a product which must meet a set of quality criteria. Viegas et al (Viegas et al., 2004; Viegas et al., 2007) found from their census of all entries existing on Wikipedia as of October 2005 that 14.5% had associated discussion pages. The average number of edits per page at that time was 15 with a median of just 2. Pages which had a high number of edits were much more likely to have discussion pages (94% of pages which had more than 100 edits had them). Through their analysis of discussion pages (Stvilia et al., 2005) identify the following Wikipedia roles which they attribute to agents:

1. Editor agents/role
2. Information quality assurance agents/role
3. Malicious agents
4. Environmental agents

The fourth type of agent is a representation of real world states which may change and in so doing render articles out of date. The main group of agents under number two are agents with administrative privileges (some 431 in April 2005). These roles may have a many to one link to actual people.

The **Meta-wiki** was established in 1994 to support discussion of community principles and issues of direction. It was established to take this form of discussion off of the entry pages and their associated discussion pages and to provide a separate forum for the debate of issues and policies affecting the whole community and the quality and direction of the Encyclopaedia itself.

**Mailing lists** provide a forum for more targeted discussion of issues, including how to address problems or disputes that have not been dealt with through local debate. These forums support mediation and more directive intervention where necessary.

Scope for formal sanction is very limited: It exists in the form of the Advisory Board and the Arbitration Committee and by the means for community members to recommend members to these bodies by vote. Persistent ‘problem people’ can be excluded from the community (i.e. have access blocked). Responsible contributors can also be allocated privileges by these committees giving them authority to place limited or complete restrictions on changes to pages for set periods of time (these people adopt the roles of ‘administrator’ or ‘bureaucrat’ depending on the level of control they are granted) .

No formal support is provided in the Wikipedia platform to assist with accounting for reputation other than in the history list traces of contributions made by different individuals. Reputation can play an informal role as administrators, editors and users observe the behaviour of contributors as recorded on these lists and the quality of their content contributions and discussions on the discussion pages. These observations may then influence their nomination and voting patterns. Reputation can also be used by the committees to determine disputes. While low key and far from the formalised reputation scoring systems used in some other sites the top 1000 contributors are identified by user name on Wikipedia and it may be possible to identify these people by name (i.e. through their personal web-page) and there have been instances of such people being named in media reports.

As reputation cannot be attained unless the editor is a registered user, Anthony et al (2005) argue that the behaviour of registered compared to non-registered contributors may be one way to

ascertain the relative influence of reputation and group identity in motivating involvement. Contributors who are motivated by reputation will register while those committed to the community will be high contributors but may or may not also be registered. Registered users with a high level of contribution are arguably motivated by both. Those that contribute seldom (perhaps only once) and anonymously (i.e. are unregistered) could be vandals, alternatively they may be ‘good Samaritans’ (Anthony et al 2005).

Those who make limited contributions anonymously are clearly not motivated by reputation and nor are they part of the community. They may, however, be committed to their own field or area of interest and hence only contribute to articles related to that field. As registration would increase the cost (time) of making a contribution, they rationally keep this cost to a minimum while nevertheless making a constructive and meaningful contribution.

Anthony et al establish a number of hypotheses around these themes and test them against data taken from the Dutch and French language sites (introducing a cultural dimension which may be of significance as there is some evidence of different behaviour on the basis of national culture on Wikipedia ). The dependent variable used is a quality proxy derived from the percentage of the contributors edit that remains in the current version of the article – i.e. its ‘survivability’. Their analysis revealed that overall, registered users contributed more content more often compared to unregistered. However, measured using the ‘survivability’ index, anonymous users contributed higher quality content. Good Samaritans contributions overall proved to survive longest. This finding is difficult to explain from a social science perspective, let alone a rational actor one. The attributes of Wikipedia – in particular its low cost of contribution as well as the large critical mass it attracts may go part way to explaining it. Anthony et al conclude ‘...open source production enables the exploitation of untapped productive resources that overcome barriers to efficient production of collective goods.’ (2005: 20). This position tends to support those who argue against the use of rules and expert criteria as any level of increased cost will reduce the capacity of these systems to scavenge valuable input from highly diffused resources. Those who argue for complete openness may, therefore, find some evidential support in this finding.

**Featured articles** provide examples and therefore a means of reinforcing the standards without recourse to complex rules.

Wikipedia moved from the only rule being ‘there are no rules’ in the early stages to now having a lengthy list of rules, procedural guides and protocols. The majority of these have emerged from the community as a means of reinforcing or encouraging (as few means of direct sanction exist) appropriate behaviour. The technology supports their use in allowing easy reference and linking to the rules within discussions.

### **3.1.2.2 Wikipedia culture**

#### **3.1.2.2.1 Beginnings**

In the beginning, complete openness was seen as valuable to encourage all comers and to avoid potential them feeling intimidated. Radical collaboration – allowing everybody to edit everyone’s (unsigned ) articles also avoided any proprietarily orientation and its attendant defensiveness as well as removing bottle necks associated with ‘expert’ editing. That said, the handpicking of a few core people is regarded by Sanger as having had an important and positive impact on the early development of Wikipedia. Sanger argues for example ‘*I think it was essential that we began the project with a core group of intelligent good writers who understood what an encyclopaedia should look like, and who were basically decent human beings.*’(Sanger, 2005). In addition to ‘seeding’ the culture with a positive disposition, this statement highlights the potential importance of establishing a style consistent with the Encyclopaedia genre – a stylistic model which might shape the subsequent contributions of others.

He concludes: *'That's pretty much it. The focus on the encyclopaedia provided the task and the open content license provided a natural motivation: people work hard if they believe they are teaching the world stuff. Openness and ease of editing made it easy for new people to join in and get to work. Collaboration helped move work forward quickly and efficiently, and posting unedited drafts made collaboration possible. The fact that we started with a core of good people from Nupedia meant that the project could develop a functional, cooperative community. Neutrality made it easy for people to work together with relatively little conflict. And the Google effect provided a steady supply of "fresh blood"--who in turn supplied increasing amounts of content.'* (Sanger, 2005).

Wikipedia echoes principles and arguably works for the same reason that other OSS projects work (Rossi, April, 2004): in part due to it having adopted a particular value position. The free license appeals to and motivates contributions 'in the public interest' and identification with the community increases the likelihood of ongoing involvement (Rullani, 2005) as well as the likelihood of participants conforming to the 'norms' of that community (Smith et al., 2007).

### 3.1.2.2.2 *Rules and norms*

There is growing evidence that the different language Wikipedia are developing somewhat distinct norms and structures (see for example some comparative statements in Riehle 2006). The following account concentrates primarily on the English version, as most research concentrates on this. The preceding technical characteristics have both shaped and been shaped by the emerging Wikipedia culture. Sanger states, *'...what we did not have worked out in advance was how the community should be organized, and (not surprisingly) that turned out to be the thorniest problem. But the fact that the project started with these good people, and that we were able to adopt, explain, and promote good habits and policies to newer people, partly accounts for why the project was able to develop a robust, functional community and eventually to succeed.'* (Sanger, 2007a).

In the early stages Sanger argues the need was for participants more than rules and so the only rule was 'there is no rule'. The reason for this, he explains, was that they needed to establish experience of how wikis should work, before over prescribing the mechanisms. However, he goes on the note *'As the project grew and the requirements of its success became increasingly obvious, I became ambivalent about this particular "rule" and then rejected it altogether... I always thought of the rule as being a temporary and humorous injunction to participants to add content rather than be distracted by (then) relatively inconsequential issues about how exactly articles should be formatted, etc.'* (Sanger, 2007a).

Very quickly, however, in the minds of some members of the community, it had become 'the essence' of Wikipedia.'

Similarly Sanger states *'deference to expertise was a policy that at least I usually insisted upon, but not strongly or clearly enough. It was nearly a year after the project began that I finally articulated this view reasonably clearly as a policy to consider. Perhaps this was because, indeed, most users did make a practice of deferring to experts up to that time.'* (Sanger, 2007a)

Sanger argues that in the early stages 'force of personality' and 'shaming' was the only means used to control contributors and that no formal exclusion occurred for six months, despite there being difficult characters from the beginning. The aim was to live with this 'good natured anarchy' until the community itself could identify and posit a suitable rule-set – i.e. the aim was to grow the rule-set based on experience of what was needed and what might work. Sanger notes that this took place within the context of a rapidly developing wider wiki culture which was opposed to rules of any kind.

Within Wikipedia rules evolved and as new ones were needed they were added to the 'What Wikipedia is not' page. *'...actual project policy, or community habits, succeeded in being established only by being followed and supported by a majority of participants. It was then, we*

said, that there was a "rough consensus" in favour of the policy. And consensus, we said, is required for a policy actually to be considered project policy. For our purposes, a "consensus" appeared to consist of (1) widespread common practice, (2) many vocal defenders, and (3) virtually no detractors.' (Sanger, 2007a)

These essentially clarified the difference between an encyclopaedia and a dictionary for example, as well as making clear the aim of producing a credible encyclopaedia which maintained the hallmarks of the genre. This clarification of goal – identifying what Wikipedia was intended to be by comparing it with what it was not, was initially the primary means for steering contributions. Wales then added the 'Neutral Point of View' (NPOV) page which emphasised the need for contributions to be free of bias. The combination of clear purpose and the principle of neutrality provided a reference point against which all contributions could be easily judged. Sanger regards the many other rules which have evolved since as secondary and not essential for success. He argues that these initial policies were arrived at in the first nine months and that the culture had consolidated within the next nine.

Reagle (2004) casts Wikipedia fundamentally as a forum for interdependent decision making and as a space of continuing discussion. He also places considerable importance on the fact that Wikipedia has a common goal and a simple and clear guide to achieving it in the Neutral Point of View. Its egalitarianism reduces sources of conflict identified on other online environments e.g. asymmetric dependence and hierarchy. He notes that critics often focus on the contentious (anarchic) aspects of the community, but cites the low levels of formal intervention (mediation or arbitration) and the relatively small number of controversial sites (approximately 600 at the time of writing) and of vandalism as a counter to this perspective. He concedes, however, that even relatively low levels of disputation can be costly to those involved with them. Reagle argues that NPOV presents an important integration potential and that its use in the Wikipedia community, far from suggesting some unattainable goal of objectivity, '*recognizes the multitude of viewpoints and provides an epistemic stance in which they can all be recognised as instances of human knowledge*' in seeking to achieve a 'fair' representation of competing positions - '*Articles should explain without advocating, characterise without engaging, and honour the intellectual independence of the reader by refraining from dogmatism*'.

The goal of building a quality encyclopaedia contributes to the avoidance of endless entanglement and in divergence that may arise from contention and competition between irreconcilable positions. Reagle draws a connections between the rules of conduct and guideline for dealing with conflict that have emerged from the Wikipedia community and existing theory of dispute handling, noting that the recommended practices align well with principles and procedures advocated in that literature about '*...how to encourage productive interdependent collaboration*'

A distinctive characteristic of Wikipedia then, in contrast to many other online discussion forums, is that the objective of the task – to produce a credible encyclopaedia– requires that differences of opinion or conflicting evidence be resolved. This may be by weight of argument, by compromise, by presenting irresolvable arguments side by side or by avoiding the area of controversy. Viegas 2004 suggest that Wikipedia supports community introspection – observation of what each other does and discussion of difference in a context that requires some form of agreement – preferably a consensus. They point in particular to the artefacts of watch-lists, discussion pages and the rule of NPOV as central to this.

Despite this argument that a few central aspects and principles do most of the coordination work in Wikipedia, there has been a proliferation of additional rules, principles and guidelines offered to orientate contributor behaviour. Even more are embedded in quality assurance bots and style-sheets etc.

### Learning and Policing the Rules

Wikipedia is generally a consensus based community, which relies extensively on internalised rules and a sense of identity on the part of individuals to ensure that positive contributors prevail over those who seek to diminish the quality and reputation. Democratic (majoritarian) procedures are used on occasions (to take straw polls) but reasoned debate is the primary means for resolving dispute, giving way where necessary to mediation, arbitration and ultimately imposition. But how to newcomers learn the rules?

Bryant et al (2005) note that there is evidence of ‘legitimate peripheral practice’ within Wikipedia. Legitimate peripheral practice is the process observed in many communities of practice whereby newcomers learn the rules, norms and skills of participation, by serving a kind of apprenticeship. They begin by undertaking minor and bounded tasks of relatively minor consequence, and then, as they become more skilled and once they have internalised the relevant knowledge and principles, gradually progress to more advanced tasks and significant roles. These authors argue that this is evident in new editors of Wikipedia initially undertaking minor editing tasks on a topic of particular interest, then moving to more significant contributions, and possibly, eventually, taking administrative roles. These authors tend to project a rather idealistic view of involvement, however, overlooking a key attribute of the Wiki environment - the lack of constraint imposed means that newcomers have the same rights as long standing participants, and as experts and therefore this mechanism for socialising newcomers can be effectively bypassed. The only restriction which they do note is a selective awareness of the range of tools available, such as mailing lists, page histories etc. Again though, these are well documented on the site and readily available for those who care to look. Careful newcomers may choose to involve themselves gradually while they ‘learn the ropes’ but this is in no way imposed by the wiki environment. These findings are also to be expected given the methodology used. This very much favoured selection of responsible contributors as they were a) registered users, and b) had an average involvement time of 14 months. They also volunteered to participate; something a ‘troll’ would be reluctant to do. Nevertheless, this research gives insight into the motives and behaviours of responsible users. It traces the process of becoming involved in a community as opposed to simply doing a task. They note for example: ‘...when asked why they contribute to the Wikipedia, many Wikipedians recognize the projects overarching goals, the appeal of community, and perceived contributions to society’. (Bryant et al., 2005).

Further research into this issue of motivation would be of value as the behaviour of wilfully destructive or naïve and un-socialised users can drive away valuable contributors.

The deep commitment evinced by some wikipedians has parallels in the wider OSS community. This phenomena in OSS is often explained (for an excellent overview of the various arguments see Rossi, April, 2004) by arguing for reputation. In OSS it is possible to gain reputation which may be usable in the wider world. In the Wikipedia environment, while some recognition of past conduct may come in the very rare event of the editor being elected to some office, this is very unlikely. Unlike OSS, there is no list of contributors to which an editor can point as evidence of their hand in the project (although they can self-identify their contributions on their user page) . Contributions are, in essence, non attributable. Identity, with product, community and values, then appears as a more likely reason for remaining involved.

### Conflict

How effective are the rules and what is the cost of maintaining them? In a study specifically designed to study the conflict and coordination costs of Wikipedia (Kittur et al., 2007: 453) note that there has been a significant increase in such costs over time. ‘...direct work on articles is decreasing, while indirect work such as discussions, procedure, user coordination, and maintenance activity (such as reverts and anti-vandalism) is increasing’. These authors used a learning algorithm targeted at pages tagged as controversial to identify metrics which signify controversy. The most important of these was the level of revision. Anonymity played an interesting

role – the anonymity of contributors to the talk page correlated with conflict while anonymity on the article itself was negatively correlated with conflict. They note that the proportion of edits made on articles has fallen from 90% in 2001 to 70% in July 2006. The proportion of indirect edits (ie those on discussion or support pages) has increased from 2% to 12%. Kittur et al cite an interview respondent as stating *‘the degree of success that one meets in dealing with conflicts (especially conflicts with experienced editors) often depends on the efficiency with which one can quote policy and precedent.’* (Kittur et al., 2007: 454). This suggests that force of argument supported by the existence of the formal rules and etiquette are important to the governance process.

Buriol et al (2006) also note that the incidence of reverts has grown over time, to stand at about 6% of edits in January 2006. Their research suggests that the three revert rule (no user to revert more than three times in 24 hours) was effective - immediately decreasing the percentage of double reverts. Overall however, there remains relatively little research into the role and effectiveness of rules and of wider normative mechanisms to the self-regulation of Wikipedia.

### 3.1.2.3 Quality Issues

#### 3.1.2.3.1 What are the measures of quality?

Stvilia et al (2005: 7) state: *‘Quality is often defined as ‘fitness for use’. It is ultimately a social construct.’* Not surprisingly then a range of different approaches have been adopted to evaluating the quality of entries on Wikipedia.

Emigh et al.(2005) evaluate the quality of Wikipedia entries by comparing them to characteristics of the encyclopaedia genre. They measured the similarity between the presentation and style of Wikipedia compared to well regarded and established print based equivalents. These authors conclude that the goal of producing a quality encyclopaedia was largely achieved when assessed by this criterion. They state: *‘Surprisingly, Wikipedia is statistically indistinguishable from the print encyclopaedia in terms of the formality features measured in this study’* (Emigh & Herring, 2005). Asking how an open environment such as the wiki can ‘reproduce traditional print norms?’ they conclude that editors have internalised the criteria of what an ‘encyclopaedia should be like’ from the wider culture. In a similar approach Elia (Elia, n.d) performed lexical analysis on both Wikipedia and Britannica. The latter tended to have more formal but shorter articles but each had similar average word length. Again the conclusion was that Wikipedia was close to taking the form and matching key characteristics of the genre.

Several assessments have been made of the factual accuracy of Wikipedia entries. The most frequently cited is a study conducted by the journal Nature in December 2005 (Giles, 2005). This was claimed to be the first evaluation to use blind peer review to assess comparable science related entries in both Wikipedia and Britannica. This revealed numerous errors in both but no significant overall difference in accuracy between the two. Of 50 articles sent to review, 42 usable reviews were returned. These revealed a total of four serious errors of fact in each and many more minor errors, omissions or misleading statements totalling 162 in Wikipedia and 123 in Britannica. There were some additional concerns about readability by some reviewers of Wikipedia articles. However, one criterion not considered was currency – an area where Wikipedia should have had an advantage.

While these comparisons for concordance with key criteria of the genre and or direct assessments of relative quality where the same or similar subjects are carried in alternative encyclopaedias are quite justified, not least due to the stated aims of Wikipedia itself their relevance have been questioned. Meyer (2006) for example argues that while individual entries in Wikipedia may well match what the best edited encyclopaedias can achieve, there will inevitably be entries which fall short, either through lack of attention – being in the process of becoming – or due to vandalism. He argues that Wikipedia is not in reality competing with peer reviewed sources but rather with alternative on-line

resources, a view supported by Magnus (2006). He argues that the community based discipline to which Wikipedia articles are subjected, means that it can often be relied upon to present more balanced information than many of the alternatives. Nevertheless, there is no call to relax the need to be a critical information user.

It would seem justified to pose a more challenging question of these comparisons on the basis of genre. What are the defining assumptions of the Encyclopaedia genre and are they defensible? The traditional printed Encyclopaedia is very much a product of the Enlightenment. It embodies the assumptions about knowledge typical of that period: knowledge as comprised of accumulations of established fact. Accordingly, the traditional encyclopaedia was developed and revised by experts, by incremental addition or adaptation of established content. While there are facts that may be readily verified (dates of the birth and death of historical figures for example), this genre obscures the fact that most areas of knowledge are contested. The Wikipedia, by contrast, has characteristics consistent with post-modern conceptions of knowledge. Knowledge is dynamic, in a constant state of construction and contention: established facts are updated, replaced or substituted, partly on the basis of new evidence but also according to changing social beliefs and the status of alternative social narratives. Wikipedia reveals the contentious and contested nature of knowledge and through the meta-perspective provided on the discussion pages, allows users insight into the nature, origins and implications of contemporary debate about issues. None of these insights are possible with the traditional genre. The question arises therefore; should different quality criteria be used to assess Wikipedia – indeed is the Wikipedia representative of a new genre defined by post-modern conceptions of knowledge and does it therefore stand, not so much in comparison to traditional forms but in contrast to them? Wikipedia arguably encourages users to be critical contributors not passive consumers of knowledge.

### Proxy Measures

In addition to these formal assessments of quality, a variety of indicators have been proposed. Andrew Lih (2004) suggests two metrics as proxy indicators of quality. The first of these is **rigor**. Rigor is measured by the total number of edits undertaken on an article on the assumption that more edits implies a deeper treatment of the topic and greater attention to its presentation and readability. The second is **diversity**, measured by the total number of unique editors. More editors are assumed to imply a wider range of perspectives contributing to depth and breadth of coverage and hence quality. These assumptions were not, however, backed by qualitative evaluation of the entries by Lih himself.

Along similar lines, Stvilia et al (2005) argue that a useful set of ‘home grown’ quality criteria exist in the form of the criteria for Featured Articles (i.e. articles voted by the community as being of high quality). The authors systematically compared these with criteria used in two relevant models drawn from the wider literature on data quality – specifically those developed by Crawford and Gasser and Stvilia. Stvilia et al then analyse the negotiations which take place among editors (on the discussion pages) as a means for identifying the types of information quality issues occurring in Wikipedia, along with their attributed causes and proposed solutions. They note that a distinguishing difference between featured articles and articles chosen at random is that with the former there generally exists ‘...a small core group of editors which is relatively homogeneous in terms of sharing social norms of cooperation, including communication protocols.’ (2005: 13). They noted that the discussion pages on featured articles were more than 10 times longer on average as well as being better organised and more readable. They argue that ‘Having well developed work articulation artefacts in the form of discussion pages helps in establishing a sense of community and negotiating a merit based social order. It helps to establish norms and conventions of communication and to introduce newcomers to those norms and the subject in general.’ (2005: 14). They note also that voting and polling proved to be effective at resolving disputes in these sites. In a separate work (Stvilia et al n.d) the authors derive seven metrics which they argue are relevant to Wikipedia quality. These are:

- Authority/reputation of editor
- Completeness
- Complexity (readability)
- Informativeness
- Consistency
- Currency
- Volatility

These were derived from factor and cluster analysis performed on article profiles to reduce an initial 19 quality metrics to a set of usable and related measures. These resulting metrics were then tested and were found effectively to discriminate between the Featured Articles included in their original sample and other randomly selected articles.

Time is a noted requirement for an article to improve in quality. Sanger refers to this as do Stvilia et al– the latter noting that Featured Articles, in addition to being 18 times larger than randomly selected articles, were also on average 3 times older (1153 days compared to 385) and thus subject to many more edits (257 compared to 8). This metric is consistent with the Rigor metric proposed by Lih (2004). The relationship between quality and time was tested by Wilkinson and Huberman (2007: 11) who conclude that ‘...a small number of articles, corresponding to topics of high relevance or visibility, accrete a disproportionately large number of edits’ and that ‘Wikipedia article quality continues to increase, on average, as the number of collaborators and the number of edits increases’. The need to include visibility in the analysis was supported by Lih (2004) who found an increase in quality after an entry had been cited in the mainstream press. It can reasonably be concluded then that most effort goes towards improving those articles which are most used. The implication of these findings is that articles which attract little interest may persist for extended periods at a low state of development and may therefore be of poorer quality. This problem is partly addressed within Wikipedia by the capacity for editors to flag articles as ‘NPOV’, ‘disputed’ or ‘accuracy disputed’ etc. thereby focusing greater attention on them. The recent incorporation of ‘notrace’ to limit search engine indexing for a period of time after an article is first submitted allows for it to accumulate edits but at the same time restricts the attention it may gain, thereby potentially slowing its rate of improvement.

### **3.1.2.3.2 Cultural contribution to quality**

A remarkable aspect of the Wikipedia environment is that many editors demonstrate a high level of commitment to, and felt responsibility for, the quality and integrity of the project (Bryant et al., 2005). This must be seen to come from the community’s values as it is not enhanced in any way by the design of the wiki which, in supporting radical collaboration, means that all contributions are non-attributable. Nevertheless many contributors maintain an interest in the integrity of articles, to which they may have made a contribution, but which are most likely to be a collective product.

Viagas et al (2007) provide evidence that Wikipedia users’ progress to take increased responsibility for more site wide coordination and quality tasks over time. This evidence was obtained by comparing the relative growth in activity in name spaces concerned with coordination and administration when compared to activity growth within the main name space of the dictionary entries themselves. This suggests that as Wikipedia complexity has grown, so too has the level of volunteer activity directed at managing that complexity.

### **3.1.2.3.3 Vandalism**

Vandalism is a particular and persistent quality problem for Wikipedia. Viegas et al (2004) developed a data analysis technique for examining interactions in a wiki space. This ‘history flow’ technique takes data from the history pages of a chosen Wikipedia article. These pages log all changes made to a page (insertions, deletions, edits etc) and who made those changes. History flow

presents transactions in a graphical way and supports statistical analysis of those transactions. Viegas et al (2004) initially examined 70 page histories compiled during May 2003. Fifty of these had long revision histories and most had suffered at least one act of vandalism. Vandalism events were categorised as:

- Mass deletion
- Offensive copy
- Phoney copy
- Phoney redirection
- Idiosyncratic copy

Analysis of mass deletions (of which there were 3574) revealed a mean time to correction of 7.7 days but a median of just 2.8 minutes i.e. 50% of mass deletions were repaired in less than 3 minutes. Obscene entries (of which there were 47) had a mean time to correction of 1.8 days but a median of 1.7 minutes. At that stage, the watch-list system was clearly sufficient to support rapid repair. Wikipedia of course grew rapidly after this period and the growing attention given to it on the web itself (the prominence of Wikipedia entries in searches) and in the mainstream media, contributed to a growing community. In 2004 a system was introduced to allow complete protection to pages which were the target of persistent vandalism or were the subject of edit wars. At the time of Viegas et al (2007) second analysis only .09% of pages were so protected. Subsequently, in 2006 a limited restriction facility was also introduced. According to these authors analysis, as of 2005 the 'self-repair' aspects of Wikipedia were still functioning effectively. Here, the median time to correct a mass deletion was 2.9 minutes and of an obscene mass deletion 2 minutes. Thus despite a significant growth in the number of pages and editors, the median time to revert vandalised pages was very similar to that measured in 2003.

More recently (Kittur et al., 2007) conducted similar analysis, noting that the overall proportion of edits due to vandalism remained very low (1-2%). Over 577, 643 edits in the data set they analysed the mean survival time was 2.1 days and the median 11.3 mins.

Anonymous authors can signal either vandalism or new users who may be less familiar with Wikipedia protocols. A high incidence of anonymous edits could, therefore, signal potential quality problems. Viegas et al 2004 analysis found 31% of edits were by such authors. They found no clear relationship between anonymity and vandalism. On the contrary they found that some pages were initially contributed and subsequently maintained primarily by anonymous contributors, notably that on libertarianism. While there is a fluctuation in the size of a page, overall the trend is for pages to grow over time.

#### **3.1.2.4 Critics of Wikipedia**

Two forms of bias occur within Wikipedia. The first is bias in the handling of individual articles. This is generally managed by invoking the need for a Neutral Point of View – i.e. by soft regulation. Bias may also accrete by the selective coverage of Wikipedia as a whole. Holloway et al (2007) note for example that bias arises to some extent due to the nature of the Wikipedia community. As *'these tend to be internet users with decent knowledge of the language, enough free time to contribute and sufficient technical facility'* unsurprisingly, the articles which predominate, are most lengthy and receive most editorial attention, are those that are of interest to this community. Wikipedia has particularly good coverage of contemporary and technological subject matter for example. The Wikipedia also reflects a geographical bias – generally in favour of English speaking countries.

Sanger has become somewhat of a critic, in particular of the governance model (or lack thereof). Sanger was clearly in a near unique position in the early stages of the development of Wikipedia.

He would have been acutely aware of the problems, spending a great deal of his time dealing with the problematical aspects of growth and the most difficult people, an experience noted by Angela Beesley in an interview with Dirk Riehle (2006) as one of the less desirable and de-motivating aspects of participation. He may well feel, and for good reason, that without the need for ongoing and systematic intervention, the project would not have progressed as it did. This view is, however, disputed. Despite this, his comments highlight that the performance and standards achieved do come at a cost - often a personal one and one that will not be uniformly shared. In this case the cost is with the energy supplied by those who challenge, negotiate, fix and compensate for the destructive behaviour of others. This no doubt, contributes to frustration and probably contributes to the loss of 'good' people to the community. The fact that the community has been able to continue to attract and retain such people is, however, part of what is interesting about it. Frustration and turnover presents a cost to individuals but so far the community as a whole has proven robust to the turnover to which it contributes.

Jason Scott in a presentation to Notacon 3, on April 8, 2006 (original presentation's audio is available at: <http://www.archive.org/details/20060408-jscott-wikipedia>) presents a somewhat more feverish critique which contain a hint of conspiracy theory. His exposition reflects a set of views scattered throughout the critical community and is, therefore, worth considering. Firstly Scott draws attention to a number of apparent contradictions in the constitution and management of the Wikipedia. Prominent here is the role played by Jimmy Wales. Scott draws attention to Wales's tendency to speak for the whole Wikipedia community, noting that this 'grand standing' is inconsistent with the egalitarian values he and the community otherwise espouse. He makes claims also of abuse of power on the part of Wales noting that 'Wikipedia office' edits are Wales edits and are untouchable. He also accuses Wales of breaking the Wikipedia communities own rules about deleting content, claiming that he has 'disappeared' material – possibly in breach of Open Source licence conditions.

More fundamentally Scott argues that the Wikipedia contains no means for reconciling fundamentally incompatible views and that over time this has led to increased internal politics, where some user's ideas become privileged. He argues that discussions over such controversial issues are very inefficient and ineffective based, as they often are, on opinion in the absence of expertise.

Wikipedia problems, he argues also have wider implications – the fact that it is comprehensively indexed and hence prominent on searches means that information sourced from Wikipedia propagates across the wider internet and within the user community – the user has no warranty on the quality however. Wikipedia, he argues, is so 'sub-contracted' that there is no accountability. Errors (deliberate or otherwise) rapidly appear on other sites becoming replicating and entrenched: disconnected from the Wikipedia editing process and hence from the noted phenomena for error to reduce with level of edit.

Moving toward the more conspiratorial, Scott argues that the openness of Wikipedia makes it a 'troll' magnet and that there are increasing numbers of strategic trolls biding their time within the Wikipedia. He even suggests that many administrators are super-vandals rather like super-bugs – i.e. those that have become immune to the normal strategies used to control vandalism and are now building up credibility in the community which they will then turn against it. In other words he argues that for some, destroying Wikipedia has become a 'world of warcraft' game.

Within the wider academic community and in the media there is ongoing debate about the philosophy of Wikipedia and associated quality issues [7][8][9][10] [11] with a couple of notable errors being frequently recycled as evidence of quality problems. An example of the latter is the story of John Seigenthaler Sr., a former newspaper editor, having a Wikipedia's entry claiming he had been involved with the assassinations of John and Robert Kennedy.

Other critics oppose its philosophy and claims to be a credible encyclopaedia, notably Robert McHenry, Former Editor in Chief of the Encyclopaedia Britannica (2004), as well as representatives of extreme conservatives concerned about libertarian and liberal views they see promoted by online culture on the one hand and libertarians, concerned about control of all kinds, on the other. For some extreme examples of conservative views see for example [1] [2] [3] [4]. There is also a site which mimics Wikipedia layout titled Conservapedia, the banner of which states:

'Tired of the '**LIBERAL BIAS** every time you search on Google and a Wikipedia page appears? [Our study](#) suggests that Wikipedia is **6 times** more liberal than the American public. Now it's time for the Conservatives to get our voice out on the internet! .... We don't make false claims of neutrality, as Wikipedia does. We have certain principles that we adhere to, and we are up-front about them. Beyond that we welcome the facts.

This site is notable for the lack of content it has attracted. On the opposite end of the political spectrum see [5] and [6].

There are also forums for the disaffected or just malicious. See for example <http://chuck.mahost.org/weblog/?p=1349> and <http://wikipediareview.com/index.php?showforum=36>, as well as the web site Wikitruth ([http://www.wikitruth.info/index.php?title=Main\\_Page](http://www.wikitruth.info/index.php?title=Main_Page)), its main page includes the following statement:

Wikitruth is a website dedicated to the subject of flaws and issues with the Wikipedia, another website run by [Jimbo Wales](#) and a massive, insane army of [Wikipedians](#) that he controls with his [mind rays](#). It's very hard to really [explain](#) Wikipedia, but if you visit it, it says it wants to be "the free encyclopedia that anyone can edit". Instead, however, it is often filled with [crazy people](#), experiences some issues with [manipulative personalities](#), and falls prey to [abuse](#) and [censorship](#). And that's a real [shame](#).

This is quite indicative of the quality of content contained.

### 3.2 Analysis of Wikipedia activity.

Preliminary analysis was undertaken of the Wikipedia in order to inform the development of Simulator Requirements (EMIL-S). It was intended that the results of this analysis of the empirical evidence would also be compared with EMIL-M and that such a comparison could be used to refine the model as well as the methodology for the Wikipedia case and to design the methodology for the subsequent Second Life case study.

In Wikipedia there are two classes of activity:

- editing activity; and
- conversation about editing activity.

Editing activity is the primary means for creating, amending and maintaining the content of the Wikipedia. This study is not concerned with this productive activity per-se but with the self-organising and self-regulating (governance) phenomena which make it possible. Insight into this can be gained by examining the Discussion pages which accompany many of the articles rather than the articles themselves.

The activity on the Discussion pages comprises a series of ‘utterances’ or speech acts between contributors about editing activity and the quality of product. Stvilia et al analysed these pages to derive indicators of quality. They argue that *‘Having well developed work articulation artefacts in the form of discussion pages helps in establishing a sense of community and negotiating a merit based social order. It helps to establish norms and conventions of communication and to introduce newcomers to those norms and the subject in general.’* (2005: 14). On the face of it then, these pages should provide a fertile source to support analysis of how social norms operate.

#### 3.2.1.1 Assumptions and expectations

Discussion pages associated with articles identified as controversial were chosen as they represent an area of activity where the quality of relationships can be expected to be more critical to goal attainment and where social norms could be expected to play an important role in regulating behaviour and where there may also be some examples of norm innovation as the participants struggled to deal with difficult and emotionally loaded discussions.

We expected to see attempts by editors to influence the behaviour of one another through the only means available to them – communicative acts. We anticipated that these may exhibit some regularity which would allow us to examine both the range and type of events that led to the explicit invocation of rules and norms and which revealed emergent influence patterns which were themselves normative. This evidence could be used to test the model being proposed in EMIL-M.

We were interested in: a) whether existing internal or external norms were used or b) whether they emerged through a process of interaction specific to the situation. If the former, then we expected that editors would attempt to invoke Wikipedia specific rules where available, or to invoke wider social norms or ethical/moral principles where they were not. If the latter then we were interested in the role of meta-norms (principles) as a foundation for the invention of new norms appropriate to the situation. Consistent with the EMIL-M, we expected to see examples of norm; *evolution* (new norm replacing old norms; *adaptation*, modification of an existing norm to apply to new contexts and/or; *extension*, where an existing norm is applied to a different situation from that for which it was originally intended. We were interested also in the means of transmission of norms, in particular whether this was done predominantly by demonstration, explicit invocation or implicitly – by how things were said or done.

We wanted also to examine what conventions prevailed and how these compared and interacted with the goal of the community and its policies. A convention is defined here as a behavioural regularity widely observed by members of the community. Policies include explicit codes of conduct as well as guidelines (etiquettes) and principles.

### 3.2.1.2 Methodology

For the study we selected a sample of 35 discussion pages associated with articles flagged as controversial. At the time of the study (May/June 2007) there were 583 such articles. This sample involved selecting every 20<sup>th</sup> entry from a random start on the list of controversial articles. There were several categories that ended up with no representation using this approach. To ensure full representation of the wide range of topics (ranging from historical events and figures, through science topics to contemporary celebrities) one entry was selected at random from categories which had no representation. The preliminary analysis reported here is based on a sub-sample of 9 of these articles. While the methodology was specifically informed by the framework being developed as a part of EMIL-M, it was decided not to frame the research too tightly around this framework, thereby potentially precluding identification of alternative or contrary mechanisms for the operation of norms. This was considered important due to the general lack of clarity surrounding definitions, propositions about normative mechanisms and about how such processes worked within OSS projects.

#### 3.2.1.2.1 Coding

As the computer mediated nature of Wikipedia communication means that no behavioural cues (other than linguistic) are possible. Wikipedia can be viewed as an institution founded on networks of commitments established, maintained and modified, exclusively in and through linguistic exchange. It was expected then that norm emergence would be revealed through the exchange of speech acts. It was anticipated, however, that the process may involve quite subtle use of linguistic cues. The methodology focused on the illocutionary force of speech acts. In other words the concern was to identify speakers attempts at influencing one another using the pragmatic (consequential) aspects of language rather than the semantic (meaning) aspects.

The intrinsic openness and the consensual nature of its the Wikipedia community and limits to formal means of compulsion or sanction led us to expect that the majority of exchanges would take the form of what Habermas calls 'communicative acts' and/or acts of 'discourse' (Cooke, 1998; Habermas, 1976). Both share the characteristic of being bounded by normative behaviour rather than strategic. According to Habermas, strategic behaviour involves recourse to formal authority, power or coercion. The controversial character of the topics suggest the need for a blend of 'communicative acts', where listeners commonly take for granted a speakers claims; and 'discourse', where the speakers claims are subject to greater critical challenge and validation cannot be assumed.

There exist a wide variety of coding schemes for natural speech. We considered a range of these before choosing to use the Verbal Response Mode (VRM) taxonomy (Stiles, 1992). VRM has been developed over many years and used in a wide range of communication contexts. Stiles defines it as '*a conceptually based, general purpose system for coding speech acts. The taxonomic categories are mutually exclusive and they are exhaustive in the sense that every conceivable utterance can be classified.*' (Stiles, 1992: 15). The classification schema has attributes very attractive where there is a need (as here) to capture many of the subtleties of natural language use that derive from and rely on the intrinsic flexibility and ambiguity of natural language yet map them to a more formal or axiomatic system needed for computer simulation.

VRM uses three structural (rather than functional) principles to classify utterances. These are a concern with whether the speaker frames an utterance on the basis of:

4. his/her own or the others source of experience;
5. presumptions the speaker makes about the others experience (feeling, perceiving or intending); and
6. Whether the speaker presents the experience from his/her own viewpoint or a viewpoint shared or held in common with another.

Using the above principles all utterances can be assigned a unique code. This code is classificatory and does not require a judgement to be made about degree. However a gauge of illocutionary ‘force’ is available through the resulting modes. The modes are defined by the points of intersection of the three dimensional matrix resulting from application of the three principles.

**Table one: Descriptors associated with Verbal Response Modes**

<i>Mode</i>	<i>Descriptors</i>
<b>Disclosure</b>	Informative, unassuming, directive
<b>Edification</b>	Informative, unassuming, acquiescent
<b>Advisement</b>	Informative, presumptuous, directive
<b>Confirmation</b>	Informative, presumptuous, acquiescent
<b>Question</b>	Attentive, unassuming, directive
<b>Acknowledgement</b>	Attentive, unassuming, acquiescent
<b>Interpretation</b>	Attentive, presumptuous, directive
<b>Reflection</b>	Attentive, presumptuous, acquiescent

(Source: Stiles, 1992: 63)

The discussion pages were coded using VRM categories applied to both the literal and pragmatic intent.

In addition The following additional codes were applied (see appendix one for the full formal definition of the codes applied):

**Valence:** VRM does not capture whether an utterance is positive or negative or neutral. Codes were added to capture this as well as the main form of positive or negative utterance.

**Subject of communication:** VRM does not code the denotative meaning of an utterance. To support an exploration of the VRM strategies used as a basis for influence in discussions with different focal points (edits, administrative actions, sources, facts, evidence, argument) etc, codes have been added to identify the subject of the utterance.

**Explicit invocation of norms and rules:** While the VRM strategies chosen by speakers can reflect norms (a preponderance of attentive, unassuming and acquiescent utterances for example may reveal observance of a politeness norm), in the Wikipedia discussions norms and/or rules are sometimes explicitly invoked. The process and effect of explicit invocation is central to the EMIL-M. Accordingly a set of codes were included to capture where explicit invocation was used as an influence strategy, and also to label the type of norm or rule invoked as well as to record the implied deontic and the event which triggered the invocation.

**Validation:** An utterance will only have a chance of influencing a listener in an intended way if it is judged by the listener to be *felicitous* (Austin) or *valid* (Habermas). There is no way to objectively assess the validity of an utterance but it is possible to observe whether, through their response to the speaker, the receiver of a message has accepted the validity claims embedded in an utterance, questioned, ignored or rejected it.

**Registration Status:** As communicative style can be expected to vary depending on whether the speaker is a registered user of Wikipedia or unregistered, each utterance is also coded according to the user’s registration status.

**Threads:** related threads of utterances are coded as a block.

### 3.2.1.2.2 Initial Findings

Contrary to expectation, analysis of an initial subset of the sample revealed that utterances which involved a specific invocation of a norm or rule were rare (only 3.2%). Of these, 44% were generated in response to the form or presentation of an article, 28% in response to an editor’s behaviour, 22% in response to an edit action and 6% an article fact. Sixty Three percent of all invocations involved specific Wikipedia rules or guidelines rather than general norms. All rules were invoked by registered users while 33% of norms were invoked by unregistered users.

Sixty five percent of all utterances were phrased in a neutral or objective manner a further 22% were negative and 14% positive. The most common negative form was ‘dismissive’ (55% of all negative utterances), while the most common positive utterance was ‘encouraging’ (39% of all positive utterances). This suggests a convention of using neutral language. The observation that utterances were predominantly phrased in neutral terms should not be interpreted as saying anything about their factual basis. Some quite exotic ideas were expressed in neutral terms.

Twenty one percent of all positively phrased utterances were explicitly validated (accepted) and a further 56% went unquestioned. Only 3% of positive utterances were rejected and 20% ignored. It is surprising then that there was such a low usage of positive style, particularly as many of the Wikipedia etiquette guidelines encourage it. By comparison only 8% of negatively phrased utterances were accepted and 24% unquestioned with 27% explicitly rejected and a further 41% ignored. Negative behaviour was clearly punished. The majority (52%) of neutrally phrased utterances went unquestioned, 21% were ignored, 11% rejected and 16% accepted.

**Table Two: Verbal Response Mode Form by Intent – number of utterances coded**

Pragmatic Intent									
Form	Edific.	Confirm	Qn	Ackmnt	Interpn	Disclosure	Reflection	Adviset	Total
Edification	336	3	3	1	15	17	2	53	430
Confirmation	0	4	0	0	0	0	0	2	6
Question	4	0	55	0	1	2	1	18	81
Acknowledgmnt	2	2	0	49	0	0	0	0	53
Interpretation	1	0	0	0	30	0	0	4	35
Disclosure	101	2	4	3	7	106	0	53	276
Reflection	0	0	1	0	0	0	10	0	11
Advisement	0	0	0	0	0	0	0	84	84
Total	444	11	63	53	53	125	13	214	976

In table two above, the rows relate to the grammatical form of the utterance (its strict literal meaning) while the columns relate to the pragmatic intent of the utterance. Within VRM an utterance is coded twice, once to capture the form and once for the intent. A typical utterance may take one form but reflect an alternative intent – for example, the utterance ‘could you close the door?’ has the form of a question but the intent of advisement – the speaker intends and listener to close the door. The relation of form to intent is expressed, "in service of" (Stiles, 1992), in this case question in service of advisement (QA).

Edification in service of Edification (EE) is the most frequent form of utterance – 33% of all utterances were of this mode. The Edification mode is defined as deriving from the speaker’s frame of reference, making no presumption about the listener and using a neutral (objective) frame of reference shared by both speaker and listener. This mode is informative, unassuming and acquiescent. As a strategy for influencing others it reflects attempts to convince by neutral objective argument.

The second most common mode is that of Disclosure in service of Disclosure (DD). Disclosure is defined as being from the speakers experience, making no presumption, but being framed using the speakers frame of reference. This is summarised as informative, unassuming but directive. Unlike EE mode, DD mode represents an attempt by the speaker to impose or have the listener accept the speakers frame. Eleven percent of all utterances adopted this form.

The third most common mode is Disclosure in service of Edification (DE). The DE mode represents an utterance which is from the speaker’s frame of reference but as if it is neutral or from a shared frame. Ten percent of all utterances used this mode. This is a somewhat neutral mode where the speaker offers clearly labelled personal knowledge as information.

The fourth most common mode is Advisement in service of Advisement (AA). AA mode represents speech from the speaker’s experience, which makes presumptions about the listener and adopts the speaker’s frame of reference. It can be summarised as informative, presumptuous and directive. It commonly takes the form of ‘you should....’ Approximately 9% of utterances were in this mode. A

further 11% of utterances have the directive pragmatic intent of advisement masked by using a less presumptuous form – that of Edification or Disclosure.

Fifty two percent of all questions were ignored as were 42% of all interpretations.

It was apparent from both the coding and the qualitative data that a great many utterances went unchallenged (47%) or were ignored (25%). A high proportion of unchallenged utterances are expected if communicative acts were predominant. The relatively low level of explicit challenging suggests that discourse involving an exploration of meaning and intent was not as prominent as expect for a discussion about controversial topics. It was clear from the discussions that this was not due to wide acceptance of one another's utterances but rather selective engagement on particular topics.

### **3.2.1.2.3 Discussion**

What is significant about the utterance strategies is that they typically involve an exchange of assertions delivered with a neutral – i.e. non-emotive style. There are very few explicit praises, or put downs and few niceties like explicit acknowledgements of one another. Seldom do contributors refer to one another by name – the exchanges are rather impersonal. This does not tally with what one would expect if the Wikipedia etiquette (<http://en.wikipedia.org/wiki/Wikipedia:Etiquette>) had been institutionalised. If we assume that the etiquette captures the community's ideal, the emerged conventions do not conform to that 'ideal'. Similarly we see low levels of questioning or of reflection (i.e. feeding back the words of the speaker to check understanding or to come to better understand the others intentions). This is arguably inconsistent with the task needs – the need to reach consensus on controversial topics. The frequency with which utterances were ignored also suggested low engagement by participants in the discussion. Why might this be?

Becker & Mark (1997) suggested that computer mediation influences conventions by altering perceived social presence - the sense of being together - and the ease of handling functionality which allows users easily to communicate social information. The concept of social presence remains somewhat ill defined (Biocca et al., 2003; Biocca et al., 2002). Nevertheless in providing only asynchronous text based public communication that is forever 'on the record', Wikipedia can be judged to support only low levels of presence.

The observed conventions may therefore be, at least in part, due to the technical characteristics of the wiki rather than reflecting social artefacts.

The absence of any expression of intimacy or acknowledgement of emotions and/or similarity of attitude (homophily) among many contributors suggests that Wikipedia lacks many of the qualities of verbal exchange that would identify it as strong community. It is more consistent with being a place to share coordination of a task. This could suggest that the goal is the primary orientating point. However, the lack of quality of discourse needed to achieve consensus is more indicative of a brief encounter between different and established positions which struggle to find common understanding rather than of a community committed to a common goal (Becker & Mark, 1997). This might suggest that the shared goal may be subordinate to more personal goals by a considerable proportion of contributors. Or it may be that the technology and environment simply will support no more than this. This environment includes the existence of saboteurs who can use the opportunity afforded by the open and anonymous platform to use identity deception i.e. to mimic the language and style of an 'expert' or to present as a genuine editor while trying to pursue personal or political agenda hostile to the aims or interests of the Wikipedia. The discussions about controversial articles provide particularly fertile ground for such sabotage. This could have an overall influence on the type of conventions which arise. Editors may, for example, display reserve and suspicion, withholding trust and taking conventional signals of authority and identity (Donath, 1998) as unreliable. The first principle in the Wikipedia etiquette is 'assume good faith'. To do so would, however, leave the process more vulnerable to 'troll' activity. This is more suggestive of the convention having arisen as a social artefact based on what works rather than concern with 'unrealistic' explicit codes of conduct.

Utterance strategies between registered and unregistered editors did not vary greatly, although unregistered editors were more likely to use disclosure intent and more likely to ask questions (possibly associated with the increased likelihood that they are relatively new to Wikipedia). There was no significant difference between registered and unregistered editors either on the tendency to use neutral compared to positive or negative utterances. The majority (55%) of both used neutral tone (57% for unregistered and 54% for registered) with approximately equal distribution of the remainder being positive and negative.

There was considerable evidence of mind reading (theory of mind) – i.e. editors appeared to form judgements about the intent of others on relatively little information. There was, however, little evidence of the use of utterance strategies to better understand or check these theories of mind. The latter would include the use of questioning, reflection, interpretation and confirmation. Editors appeared quick to judge and to then follow response scripts consistent with those judgements (e.g. ignoring or accepting utterances of others) on the basis of those judgements. Consistent with this, there were few instances of renegotiated patterns of communication style. Positions and styles stayed relatively constant over the period of the interaction. Only occasionally would an editor modify his/her style significantly if challenged. Of the rule invocations 26% were accepted, a similar proportion were rejected or ignored and the remainder went unquestioned (but generally had no affect on behaviour). This is consistent with norms viewed as being triggered by a limited range of cues which allow individuals to locate themselves and select identities appropriate to a context and which then remain essentially stable.

#### 3.2.1.2.4 Observations

In this preliminary analysis we set out to identify how normative regulation appears to work in this environment. The findings have challenged some of our assumptions and expectations, in particular:

- The more detailed and specific behavioural etiquette seems to have little influence on the overall character and style of interaction.
- The overall quality of interaction of editors falls short of the range and quality of communicative style characteristic of a community and consistent with what one would expect given the nature of the task.
- Most regulation is achieved without the need for frequent explicit invocation of rules or norms. Rather, behaviour seems to accord to a convention or pattern which editors quickly recognise and conform to and which minimally accommodates what needs to be done to satisfy the task in a context of somewhat heterogeneous personal goals.
- There was a lack of evidence of active negotiation of expectations and standards and convergence of behaviour towards a norm. Within the discussion pages there appeared to be an accommodation of a set of conventions and little obvious norm innovation, evolution, adaptation or extension. This suggests that on first encounter with Wikipedia editors read a set of cues as to what constitutes appropriate or acceptable behaviour and then make operational a script which accommodates it and stays relatively constant.

This is perhaps consistent with what Bicchieri was advancing when she says that in activating a norm:

*...we go through a complex set of mental operations of interpreting, understanding, encoding and inferring, the output of which is meaningful, appropriate behaviour. These operation include (but are not limited to) the perception of a stimulus, be it a person or an event, focusing attention on particular cues or dimensions of it, activating a comparison process to assess similarities and differences with past episodes stored in memory, grouping the stimulus with others in a sensory category, and finally invoking a cognitive schema that will specify beliefs, expectations, and behavioural rules.'... 'Categorization is a crucial step*

*...as it activates schemata (or scripts) that may be likened to personal theories of the way social situations and people work (Bicchieri, 2006: 81)*

Some further insight into this may come from identity theory. Postmes et al (2005) expressly considers identity in terms of micro-macro interpenetration, examining the way in which individual identity shapes social identity and vice versa within small groups. He suggests two mechanisms; *categorization*, where the individual distinguishes groups and locates themselves within the space of alternative associations available in a given context, and; *internalization*, where the individual takes on the ‘stereotypes and norms’ appropriate to groups with which they identify.

Categorization is described as a deductive top-down process based on comparison. Once the agent has categorized and selected the group/s with which it identifies based on a reading of environmental cues then the perceived difference between ‘in-groups’ and ‘out-groups’ is psychologically maximised (the agent concentrating on the positive characteristics of the ‘in-group’ and negatives of the ‘out-group’). This process may be a part of resolving a boundary problem (Paulsen & Hernes, 2003) – how to determine the criteria which serve to bound one group as ‘alike’ and perhaps ‘like me’ and other groups as different in the presence of multiple dimensions, at least some of which overlap. In this way the agent reduces its individuality (agency) adopting an ‘identity’ consistent with the norms of the ‘in-group’. The identities are ‘...cognitive schemas-internally stored information and meanings serving as frameworks for interpreting experience...they increase sensitivity and receptivity to certain cues for behaviour’ (Stryker & Burke, 2000: 286). This raises some interesting questions about what influences an agent to choose among the infinite bases of possible comparison.

Agents participate in many intersecting and non-intersecting social domains, each involving the agent in different roles, so agents may develop somewhat different identities appropriate to those roles. Thus a single ‘self’ may be able to present an ‘ecology’ of identities (Smith-Lovin, 2003). Some of these roles may be more central to ‘self’ than others and within identity theory are argued to form a hierarchy with identities closer to the top more likely to be invoked. Wikipedians may identify as members of the Wikipedia community AND with members of a group which holds particular beliefs. This carries some possibility of identity conflict and this may be resolved in different ways at different times. He/she may generally follow the norms of Wikipedia unless or until he/she feels central values are being challenged and then switch to a set of norms of an alternative group.

In addition to categorisation Postmes suggests that the process of identity construction arises through ‘*explicit or implicit negotiations over different (competing) understandings*’ We expected to see evidence of this happening on the Discussion pages but did not. If this process follows categorization then perhaps is unlikely to happen on the article Discussion pages but rather to occur behind the scenes as like minded individuals (those who have categorised one another as belonging to the same ‘in-group’), get together on personal talk pages or share views by more confidential means about those they are beginning collectively to see as ‘other’.

## 4 Future steps

The research to date has raised a number of questions which require further investigation and need to be reconciled with the assumptions which have informed the EMIL-M. The case also presented limited options for understanding the intra-agent processes important to the development of the EMIL-A. Future work will therefore be directed at exploring these aspects.

### 4.1 Extended analysis of Wikipedia Case

We now propose to conduct the same analysis for both archival discussion pages (from a period before the rules of Wikipedia became established) and of current Featured Articles in Wikipedia. Our expectation is that in the former we may see more active use of norm invocation as a) editors will be more likely new to the wiki environment and b) there are few situation specific rules to draw on leaving only recourse to wider social norms as a means of checking inappropriate behaviour of

new users. Features Sites reflect sites of high quality and this may be based on effective social coordination of a diverse range of talents. It may be that greater community spirit will be evidenced on these pages.

The data available through both article pages and discussion pages does not support analysis of the internal choices or decision processes involved in norm recognition, the formation of beliefs about norms, evaluation of norm salience nor the formation of norm intention. Nor does it allow the effect of group identity and or goal hierarchy to be examined as information about individuals responses, processing and judgements are not revealed through the available evidence. In an attempt to rectify this, a trial questionnaire directed towards Wikipedia editors was developed and piloted. While the response rate for the test approximated 40%, some editors took offence at the means of delivery of the questionnaire (placing a link and background information along with an invitation to participate on their personal talk pages) even though care was taken to minimise the intrusion. A further difficulty with this approach is that the sample of Discussion Pages being analysed was taken in May 2007 and the events about which questions are asked are therefore historical. It must therefore be expected that recall bias and other sources of error would reduce the quality of data collected in this way. As an alternative an experimental design has been proposed.

## **4.2 Wiki Experimental Design**

The experiment will involve volunteers in undertaking a task in a Wiki environment of a similar nature to article editing in Wikipedia. Experimental variables to be manipulated will include:

- Group affiliation
- Personal vs collective goal orientation
- The effect of visual cues as to behavioural norms on editor communicative behaviour.

An initial design for the experiment is set out in more detail in Appendix two.

## **5 Second Life Case Study**

Second Life (SL) is a 3-D virtual world built and owned by its Residents. It went online in 2003 and currently attracts approximately 10,000 ‘residents’ from around the world. As an open virtual world Second Life supports a wide range of activity – social and professional. Unlike Wikipedia it therefore provides an opportunity for participants to identify goals and to self-organise around a wide variety of goals and tasks.

Users can create groups and environments (Sims) which others may join. Groups may be open to all comers or accessible on invitation only or to those meeting certain criteria (e.g. male or female, gay etc.). Those who form the groups can (within certain constraints) define what can be done within those groups and how things should be done – including defining rules of conduct.

Unlike Wikipedia SL involves participants in synchronous communication. Until recently this was text based with the opportunity for exchanges to be broadcast to all those nearby or directed at specific individuals. The environment now also incorporates voice communication.

Also in contrast to Wikipedia users interact in a simulated three dimensional space containing a variety of artefacts and can themselves create those artefacts.

Actors take on a visual presence through avatars which they can select and modify. This adds a higher level of ‘social presence’ compared to Wikipedia. Avatars have invented SL names but users may adopt several avatars (and hence identities) and the persons real world identity is hidden from those who interact with the user thus anonymity of the users real life identity is conserved unless he/she chooses to disclose it.

Bainbridge (2007) has argued that of this type of artificial environment, SL is particularly suitable for the conduct of social scientific enquiry due to its open task environment. It is quite possible to recruit participants for involvement in situated or experimental research in much the same manner

as would be possible in the real world. This recruitment is not constrained by geography and may involve potentially large numbers at low cost when compared to conventional research.

Within second life there is no material need for social organisation and it is almost impossible to force another user to do anything against their own wishes (except social expectations and exclusion). Communities in Second Life have grown up around:

- Role play scenarios (Dragons, Combat, Science Fiction Novels, Vampirism etc.)
- Commercial activity (Marketing, Designing and Selling, Services, etc.)
- Shared interests (Music, Building, Sexual Orientation, Political Leaning, Art, etc.)
- Specific locations have developed for these communities

They are developing new norms as well as selecting which of real life norms to keep

For the purposes of EMIL SL is attractive also as it necessarily involves participants in new social environments, ones which they may participate in co-creating and in a novel environment. SL also supports a lot of social experimentation, with participants free to experiment with different identities, affiliations and behaviours with relatively low long term social costs. It is therefore, a social environment in which it may be more likely to identify the need for norm innovation or the adaptation of norms imported from more social environments outside of the virtual world.

EMIL has a presence in SL through the Centre for Policy Modelling at Manchester University. We are therefore well placed to collect data in this environment. The methodology for Second Life will involve two parallel data collections. The first will elicit information from Sim Managers (people who have established and maintain groups or sand-boxes in SL). The second will recruit people interested in normative behaviour in SL and collect information on ‘normative events’ which they encounter in their normal participation within SL. The interviews with Sim Managers will provide a macro overview of the need for and observations of collective behaviour within particular environments. The information sought from individuals will include their construal of events and information about how they recognised, evaluated and relate norms, and form normative intent. This should provide more information on the mental path to norms explicitly modelled within EMIL-M.

### **5.1 *Sim Manager Interviews***

Approximately 10 interviews with Sim Managers will be completed. The interview process will be semi-structured with all interviewees being asked the following core questions.

- What was your attitude/approach to sim management/security at the beginning?
- How has this developed over time? What worked and did not work?
- Examples of specific events that lead you to make changes in the management of the Sim, suggestions to visitors/users, security etc.
- What is your current approach to sim management/security.
- Is this sim open to the public or is access restricted? How is this done? Who is it restricted for/against?
- Do the users/visitors of your sim form any sort of informal/formal group or community?
- If so, what is it that keeps this group together or caused it to form in the first place?
- How are members of this community recognised/recognisable to each other?
- Anything else you want to say about the issue. E.g. other comments about acceptable/unacceptable behaviour in your sim or SL, its management and occurrence?

Other information collected

- Ask for copies of any suggestions/rules for av behaviour made by the sim – are there any previous versions of this you could give us?
- Location (SURL)
- Their description of the sim
- Avatar in world name
- Age of sim
- Date of interview
- Interviewer

## **5.2 Normative Events**

Some 25 participants in SL have already volunteered to participate in this research. Each will be approached to collect data relating to one or more ‘normative events’ encountered while they go about their business within SL over a specified period (approximately 1 month). A normative event is defined as an event involving two or more people in which one explicitly or implicitly expects or demands of the other compliance with some explicit or implicit standard, norm or rule. The recruit may or may not be directly involved but must have been personal witness. The participant may be the person making the normative demand.

Participants will be asked to document the event as close as possible to the time of its occurrence. The following information will be captured on a web-site provided for the purpose.

- SL ID
- SURL – location
- Screenshot - upload
- Copy and paste relevant chat/IM dialogue (please say which)

Participants will also be asked to provide a brief narrative account in response to the following prompts:

- Can you describe the event in your own words, including your opinion of why it happened, what were the main actions taken by those involved (including your own), if any, and the outcomes for those involved.
- What did you think/feel about those involved and their actions (including your own)? If you took any action please describe your reasons for doing so.
- Can you describe who in this situation already knew each other, and the nature of these existing relationships? Contra-wise, were any of those involved newcomers/outside to the situation/place/community?
- Any other comments/information you wish to make about the event (optional)

## 6 Appendix One: Wikipedia Case Study Code Frame and Definitions

Codes are applied at three different levels:

**Utterance:** defined as ‘a simple sentence, an independent clause, a non-restrictive dependent clause, an element of a compound predicate, or a term of acknowledgement, evaluation or address’.

**Contribution:** defined as a single contribution by an individual editor.

**Thread:** A series of exchanges which represent responses to prior contributions with a related subject.

### Communicative Style

Communicative style codes are applied at the level of utterance.

The communicative style codes are intended to capture

1. The general valence of an utterance
2. The attitude or style expressed by the utterance

If an utterance is **negative**, it is coded to negative and also to the style of negative.

- Abusive: To assail with contemptuous, coarse, or insulting words;
- Aggressive: characterized by enmity or ill will, threatening
- Contemptuous: exhibiting lack of respect; rude and discourteous
- Defensive: attempting to justify or defend
- Dismissive: showing indifference or disregard; not having or showing interest

If an utterance is **neutral/objective** then it is coded to this node only.

If an utterance has a **positive** intent then it is coded to positive and also to the style of positive.

- Affirming: To support or uphold the validity of
- Apologetic: Self-deprecating; humble
- Encouraging: furnishing support and encouragement
- Placative: To allay the anger of, especially by making concessions; appease

### Editor\_Status

Editor status code is applied at the level of a contribution.

**Unregistered:** is applied if the editor has no username at the end of the contribution but that contribution terminates with only an IP address.

**Registered** is applied if the editor has a username. In this case the contribution should also be coded to the username under Editor\_ID. This may require creation of a new node of that name.

### Normative behaviour

The normative behaviour codes are applied at the level of utterance.

If, in an utterance an editor:

- Specifically invokes a rule present in a Wikipedia guideline, etiquette guide or style guide, and the Editor links the invocation to that rule (e.g. ‘at Wikipedia we always provide a source to support any fact’) then code to the associated **rule\_descriptor** (create new node where necessary).
- Specifically invokes a norm which is a) not the subject of an existing rule or b) in the invocation, the editor does not link to a specific Wikipedia practice or rule (even if one exists) but rather refers to a wider standard (e.g. ‘it is not good practice to play the man rather than the ball’), then code to **norm\_descriptor**. This may require creation of a new node with an associated descriptive name for that norm.

Code to the associated **deontic** operator

- it is impermissible that
- it is non-obligatory that
- it is obligatory that
- it is optional that

- it is permissible that

Code to the appropriate **trigger**.

- Administrative\_action
- Article\_Fact
- Article\_form\_presentation
- Edit\_action
- Person\_behaviour

### **Subject of communication**

Subject of communication is coded to the level of utterance.

This is the subject of the utterance. It may need to be inferred from the context.

- About\_Administrative\_action
- About\_Edit\_Action
- About\_Editor\_person
- About\_person\_behaviour
- Article\_Fact (this includes reference to sources).
- Article\_form\_presentation (includes presentation, structure, location, format, style etc)

### **Thread**

The thread code is applied at the thread level. Contributions linked to a single distinctive thread are coded as a block.

### **Validation**

Validation is coded at the level of utterance. To apply the code it is necessary to read ahead and to determine if the utterance was:

- **Accepted:** by being met with an acknowledgement, explicit acceptance or implicit acceptance (e.g. by taking the action asked for)
- **Rejected:** by being explicitly questioned or rejected as invalid, inappropriate or where the motive of the speaker for making the utterance is questioned.
- **Ignored:** No response discernable either explicitly (i.e. there was no explicit questioning or rejection of the utterance) or implicitly (i.e. no evidence of a requested action having been undertaken).
- **Unquestioned** where the conversation proceeds without specific reference to the utterance but it is apparent from the subsequent conversation or action that the utterance was accepted.

### **VRM**

VRM codes are applied at the level of utterance.

An utterance is coded to the appropriate **mode** descriptor twice: once for the **literal** meaning and once for the **pragmatic** intent.

## Modes

### Descriptors associated with Verbal Response Modes

<i>Mode</i>	<i>Source of Experience</i>	<i>Presumption</i>	<i>Frame of reference</i>	<i>Descriptors</i>
<b>Disclosure</b>	Speaker	Speaker	Speaker	Informative, unassuming, directive
<b>Edification</b>	Speaker	Speaker	Other	Informative, unassuming, acquiescent
<b>Advisement</b>	Speaker	Other	Speaker	Informative, presumptuous, directive
<b>Confirmation</b>	Speaker	Other	Other	Informative, presumptuous, acquiescent
<b>Question</b>	Other	Speaker	Speaker	Attentive, unassuming, directive
<b>Acknowledgement</b>	Other	Speaker	Other	Attentive, unassuming, acquiescent
<b>Interpretation</b>	Other	Other	Speaker	Attentive, presumptuous, directive
<b>Reflection</b>	Other	Other	Other	Attentive, presumptuous, acquiescent

### VRM Principles

VRM principles are used to identify the VM modes. There are three principles, source of experience, presumption and frame of reference.

#### **Source of Experience: Who's experience is the topic of utterance?**

**Speaker:** Informative: utterance based on speakers personal experience, reveals feelings or things he/she knows.

**Other:** Attentive, questioning or describing and/or, reflecting others experience, inviting contribution or keeping space open for other to comment.

#### **Presumption: Does the speaker need to presume knowledge of other?**

**Speaker:** unassuming, signals of receipt of message but not agreement/disagreement or comparison. Concerns objective information and hence no need for presumption.

**Other:** presumption made about others experience or intention – how he/she is, was, will be or should be. Also utterances that seek to guide the others behaviour (i.e. to impose a view or compel an action). Includes only utterances where the presumption is necessary to the meaning. In order to agree or disagree or compare presumption is necessary unless it is about objective information held by the speaker.

#### **Frame of Reference: Would listener have to read the speakers mind?**

**Speaker:** reference is made to own constellation of meaning – directive. Listener would need to be able to read your mind to verify. Seek to impose speakers view: Includes personal perceptions, intentions, thoughts, feelings and value judgements.

**Other:** Acquiescent: reference is made to a common or shared set of meanings. Objective – external to the speaker, placeholders are objective/neutral.  
Ask who gets to say what is true – code here if objective or other.

## 7 Appendix Two: Wiki Experiment

**Aim:** to examine the effect of

- personal goals
- social norm
- group membership/identity

on product quality and communication style on a shared wikiwiki mediated writing task.

**Task:** Prepare documentation that summarizes positions with respect to a controversial topic.

**Technology:** Wikiwiki environment with Article (product output) page and separate discussion page.

**Group sizes:** 4-7 people per group.

Well before the experiment, participants will be interviewed/surveyed to ascertain their views with respect to the topic. Individuals will then be selected and allocated to groups in order to achieve a wide representation of positions in each group.

### **Experiment.**

Immediately prior to commencing the main task, participants will perform a scrambled sentence task which primes them to either a) predispose them towards favouring personal goals over collective or b) favour collective goals over personal.

They will also be told that they have been allocated to a group which is a) made up of like minded individuals like them who want to collaborate b) made up of individuals unlike them who may be difficult.

Half of the groups will then sit down at a wiki to perform the task but will view prior discussion fragments which are a) positive/friendly b) negative/combative

At the end of the task, participants will answer a questionnaire which asks about goals, norms, group identification, and assessment of product, view on what factor most influenced product quality and what could have been done to improve the task effectiveness.

This implies the need for 8 groups with each group taking up a unique combination of the three dichotomous variables identified above.

The resulting product will be analyzed from a quality perspective. The discussions will be analyzed for communicative style as per the Wikipedia analysis.

The intention is to do a test to see in what way individual vs collective goals interact with pro-social vs antisocial norms associated with group identity and/or visual cues as to past group behaviour.

## 8 References

### 8.1 Media References

- [1] 'Lib-pedia'? Anti-conservative bias rampant at wikipedia  
<http://www.themediareport.com/halloffame.htm>
- [2] U.S. senator: It's time to ban Wikipedia in schools, libraries  
 By Preston Gralla on Wed, 02/14/2007 - 7:08pm <http://www.computerworld.com/blogs/node/4598>
- [3] Wikipedia: A Nightmare Of Libel and Slander, By Joel Leyden  
 Israel News Agency  
<file:///C:/data/surrey/Case%20Studies/wikipedia/wikipedia/libelslandersexwoolencyclopedia48330508.html>
- [4] Does Wikipedia Support Terrorism? By Joel Leyden,  
<http://www.israelnewsagency.com/wikipediaterrorismiranrussialeninisraelcensorship4877031407.html>
- [5] Banned indefinitely, censorship on Wikipedia,  
[http://earthopenetwork.net/Banned\\_Indefinitely\\_Censorship\\_on\\_Wikipedia.htm](http://earthopenetwork.net/Banned_Indefinitely_Censorship_on_Wikipedia.htm)
- [6] Latest on Wikipedia <http://chuck.mahost.org/weblog/?p=1349>
- [7] Are you thinking what I'm thinking?, The Times, October 13 2006,  
<file:///C:/data/surrey/Case%20Studies/wikipedia/are%20you%20thinking%20what%20im%20thinking.htm>
- [8] Wikipedia: "A Work in Progress", Business Week , December 14 2005,  
<file:///C:/data/surrey/Case%20Studies/wikipedia/business%20week%20wikipedia%20work%20in%20progress.htm>
- [9] Insider Editing at Wikipedia, New York Times, by Dan Mitchell December 24, 2005
- [10] Knowledge to the People, interview with Jimmy Wales, New Scientist, January 31 2007,  
<http://www.newscientist.com/channel/opinion/mg19325896.300>
- [11] Can Wikipedia Survive its Own success, Knowledge@wharton, January 25 2006  
<http://knowledge.wharton.upenn.edu>

### 8.2 Academic References

- Ablowitz, R. 1939. The Theory of Emergence. *Philosophy of Science*, 6(1): 16.
- Ajzen, I. 1985. From Intentions to Actions: A Theory of Planned Behavior. In J. Kuhl (Ed.), *Action Control: From Cognition to Behavior*. New York: Springer Verlag ISBN 038713445X.
- Almeida, R. B., Mozafari, B., & Cho, J. 2007. On the Evolution of Wikipedia, *International Conference on Weblogs and Social Media* Boulder Colorado.
- Anthony, D., Smith, S. W., & Williamson, T. 2005. Explaining Quality in Internet Collective Goods: Zealots and Good Samaritans in the case of Wikipedia. Hanover: Department of Sociology, Dartmouth College.
- Archer, M. 1998. Realism in the Social Sciences. In M. Archer & R. Bhaskar & A. Collier & T. Lawson & A. Norrie (Eds.), *Critical Realism: Essential Readings*. London: Routledge.
- Archer, M., Bhaskar, R., Ciollier, A., Lawson, T., & Norrie, A. 1998. *Critical Realism: Essential Readings*. London: Routledge. ISBN: 0-415-19632-9
- Armitage, C. J. & Conner, M. 2001. Efficacy of the Theory of Planned Behaviour: A Meta analytic review. *British Journal of Social Psychology*, 40, ISSN 0144-6665: 471-499.
- Austin, J. 1962. *How to do things with words*. Oxford: Oxford University Press.
- Bagozzi, R. P., Baumgartner, J., & Yi, Y. 1989. An Investigation into the role of intentions as mediators of the attitude behavior relationship. *Journal of Economic Psychology*, 10, ISSN 0167-4870: 35-62.
- Bagozzi, R. P. 1992. The Self-Regulation of Attitudes, Intentions and Behavior. *Social Psychology Quarterly*, 55(2), ISSN 0190-2725: 178-204.
- Bainbridge, W. S. 2007. The Scientific Research Potential of Virtual Worlds. *Science*, 317: 472-476.
- Becker, B. & Mark, G. 1997. *Constructing Social Systems through Computer Mediated Communication*. Sankt Augustin, Germany: German National Research Center for Information Technology.
- Berger, P. L. & Luckman, T. 1972. *The Social Construction of Reality*: Penguin. ISBN: 0140600019
- Bergquist, M. & Ljungberg, J. 2001. The power of gifts: organizing social relationships in open source communities. *Info Systems*, 11: 305-320.

- Bertalanffy, L. v. 1950. An Outline of General Systems Theory. British Journal for the Philosophy of Science, 1(2).
- Bertalanffy\_von, L. 1968. General Systems Theory. New York: Braziller.
- Bhaskar, R. 1997. A Realist Theory of Science. London: Verso.
- Bhaskar, R. 1998. The Possibility of Naturalism. London: Routledge. ISBN: 0-415-19874-7
- Bicchieri, C. 2006. The Grammar of Society. Cambridge: Cambridge University Press. ISBN 0521574900
- Biocca, F., Harms, C., & Gregg, J.; The Networked Minds Measure of Social Presence: Pilot Test of the Factor Structure and Concurrent Validity, [www.mindlab.org](http://www.mindlab.org)
- Biocca, F., Harms, C., & Burgoon, J. K. 2003. Towards a More Robust Theory and Measure of Social Presence: Review and Suggested Criteria. Presence, 12(5): 24.
- Bowles, S. & Gintis, H. 2003. The Origins of Human Cooperation. In P. Hammerstein (Ed.), The Genetic and Cultural Origins of Cooperation. Cambridge, MA: MIT Press.
- Bruner, J. 1991. The Narrative Construction of Reality. Critical Enquiry, 18(1): 20.
- Bryant, S. L., Forte, A., & Bruckman, A. 2005. Becoming Wikipedian: Transformation of Participation in a Collaborative Online Encyclopedia, GROUP 05. Sanibel Island Florida USA.
- Buriol, L. S., Castillo, C., Donato, D., Leonardi, S., & Millozzi, S. 2006. Temporal Evolution of the ikiagraph. Paper presented at the Web Intelligence Conference, Hong Kong.
- Burrell, G. & Morgan, G. 1994. Sociological Paradigms and Organisational Analysis. London: Virago. ISBN: 1-85742-114-0
- Castelfranchi, C. 1998a. Through The Minds of the Agents. Journal of Artificial Societies and Social Simulation, 1(1).
- Castelfranchi, C. 1998b. Simulating with Cognitive Agents: The Importance of Cognitive Emergence. In J. S. Sichman & R. Conte & N. Gilbert (Eds.), Lecture Notes in Artificial Intelligence. Berlin: Springer Verlag ISBN 3-540-65476-3.
- Chalmers, D. J. 2006. Strong and Weak Emergence. Canberra: Research School of Social Sciences, Australian National University.
- Chance, T. August, 2005. The Hacker Ethic and Meaningful Work: 38: University of Reading.
- Checkland, P. 1988. Systems Thinking Systems Practice. G.B.: John Wiley.
- Christley, S., Xu, J., Gao, Y., & Madey, G. 2004. Public Goods Theory of the Open Source Development Community. Paper presented at the Agent2004, Chicago, IL.
- Ciffolilli, A.; Phantom Authority, self-selective recruitment and retention of members in virtual communities: The case of Wikipedia; 12, 8, [http://firstmonday.org/issues/issue8\\_12/ciffolilli/index.html](http://firstmonday.org/issues/issue8_12/ciffolilli/index.html)
- Clayton, P. 2006. Conceptual Foundations of Emergence Theory. In P. Clayton & P. Davies (Eds.), The re-Emergence of Emergence: The Emergentist Hypothesis from Science to Religion. Oxford: Oxford University Press.
- Clayton, P. & Davies, P. 2006. The Re-Emergence of Emergence: The Emergentist Hypothesis from Science to Religion. Oxford: Oxford University Press. ISBN: 0199287147
- Coase, R. H. 1993. The Nature of The Firm. In O. E. Williamson & S. G. Winter (Eds.), The Nature of the Firm: Origins, Evolution and Development. N.Y.: Oxford University Press.
- Coleman, J. S. 1994. Foundations of Social Theory. Cambridge: Belknap. ISBN 0674312260
- Cooke, M. (Ed.). 1998. On the pragmatics of communication: Jurgen Habermas. Cambridge, MA: MIT Press.
- Dalle, J.-M., David, P. A., Ghosh, R. A., & Wolak, F. A. 2004. Free & Open Source Software Creation and the 'Economy of Regard', Third EPIP Workshop "What Motivates Inventors to Invent?" Pisa Italy.
- Dautenhahn, K. 2002. The Origins of narrative: In search of the transactional format of narratives in human and social animals. International Journal of Cognition and Technology, 1(1): 26.
- Davies, P. 2006. The Physics of Downward Causation. In P. Clayton & P. Davies (Eds.), The Re-Emergence of Emergence: The Emergentist Hypothesis from Science to Religion. Oxford: Oxford University Press.
- Demil, B. & Lecocq, X. 2003. Neither market or hierarchy or network: The emerging bazaar governance: 36: Université Lille/Institut d'Administration des Entreprises.
- DeSanctis, G. & Poole, M. S. 1994. Capturing the Complexity in Advanced Technology Use: Adaptive Structuration Theory. Organization Science, 5(2): 26.
- Dholakia, U. M., Bagozzi, R. P., & Pearo, L. K. 2004. A social influence model of consumer participation in network and small group based virtual communities. International Journal of Research in Marketing, 21: 241-263.

- Disabatino, J.; Free Britannica bites the dust, <http://www.library.yale.edu/~llicense/ListArchives/0103/msg00049.html>
- Donath, J. S. 1998. Identity and deception in the virtual community. In P. Kollock & M. Smith (Eds.), Communities in Cyberspace. London: Routledge ISBN 0415191408.
- Edwards, K. 2001. Epistemic Communities, Situated Learning and Open Source Software Development, Epistemic Cultures and Practice of Interdisciplinarity. Trondheim.
- Elia, A. n.d. An analysis of Wikipedia digital writing. Napoli: Dipartimento di Scienze Statistiche, Università degli Studi di Napoli Federico II, Napoli Italy.
- Elliott, M. & Scacchi, W. May, 2004. Mobilization of Software Developers: The Free Software Movement: University of California, Irvine.
- Ellis, G. F. R. 2006. On the Nature of Emergent Reality. In P. Clayton & P. Davies (Eds.), The Re-Emergence of Emergence: The Emergentist Hypothesis from Science to Religion. Oxford: Oxford University Press ISBN: 0199287147.
- Emigh, W. & Herring, S. C. 2005. Collaborative Authoring on the Web: A Genre Analysis of Online Encyclopaedia. Paper presented at the 38th Hawaii International Conference on System Sciences, Hawaii.
- Engels, F. 1934. Dialectics of Nature. Moscow: Progress Publishers.
- Engestrom, Y., Miettinen, R., & Punamaki, R.-L. 1999. Perspectives on Activity Theory. New York: Cambridge University Press.
- Eriksson, O. 1999. A Generic Communications Model Based on Habermas' and Searle's Versions of Speech Act Theory. Borlange, Sweden: Dalarna University.
- Eronen, M. 2004. Emergence in the Philosophy of Mind. University of Helsinki, Helsinki.
- Escher, T. 2004. Political Motives of Developers for Collaboration on GNU/Linux. University of Leicester & Freie University Berlin.
- Evans, P. & Wurster, T. S. 2000. Blown to Bits: How the New Economics of Information Transforms Strategy: Harvard Business School Press.
- Fehr, E., Fischbacher, U., & Gächter, S. 2002. Strong Reciprocity, Human Cooperation and the Enforcement of Social Norms. Human Nature, 13: 1-25.
- Fishbein, M. & Ajzen, I. 1975. Belief, Attitude, Intention and behavior: An Introduction to Theory and Research. MA: Addison-Wesley. ISBN 0201020890
- Fjeld, M., Lauche, K., Bichsel, M., Voorhorst, F., Krueger, H., & Rauterberg, M. 2002. Physical and Virtual tools: Activity Theory Applied to the Design of Groupware. Computer Supported Cooperative Work, 11: 27.
- Fodor, J. A. 1974. Special; Sciences or The Disunity of Science as a Working Hypothesis. Synthese, 28, ISSN: 1573-0964: 18.
- Ford, P. & Dominey, A. R. n.d. Putting all the Strands Together, Co-evolution of language and theory of mind, Interdisciplines.
- Franke, N. & von Hippel, E. 2003. Satisfying Heterogeneous User Needs via Innovation Toolkits: The Case of Apache Security Software. Research Policy 32(7): 1199-1215.
- Fuchs, C. & Hofkirchner, W. 2005. The Dialectic of Bottom-up and Top-down Emergence in Social Systems. tripleC 1(1), ISSN 1726-670X: 22.
- Gibbs, J. P. 1981. Norms, Deviance and social control: Conceptual matters. New York: Elsevier. ISBN 0444015515
- Giddens, A. 1984. The Constitution of society: Outline of the theory of structuration. Berkeley: University of California Press.
- Gilbert, N. 1995. Emergence in Social Simulation. In N. Gilbert & R. Conte (Eds.), Artificial Societies. London: UCL Press ISBN 1-85728-305-8.
- Gilbert, N. 2002. Varieties of Emergence. Paper presented at the Social Agents: Ecology, Exchange, and Evolution Conference Chicago.
- Giles, J.; Internet Encyclopaedias go head to head, <http://www.nature.com/news/2005/051212/full/438900a.html>
- Goldspink, C. & Kay, R. 2007. Social Emergence: Distinguishing Reflexive and Non-reflexive Modes Paper presented at the AAAI Fall Symposium Washington.
- Gotham, K. F. & Staples, W. G. 1996. Narrative Analysis and the New Historical Sociology. The Sociological Quarterly, 37(3): 20.

- Habermas, J. 1976. Some Distinctions in Universal Pragmatics: A working paper. Theory and Society, 3(2): 12.
- Hars, A. & Ou, S. 2002. Working for Free? Motivations for Participating in Open-Source Projects. International Journal of Electronic Commerce, 6(3): 25-39.
- Hodgson, G. M. 1996. Economics and Institutions. Oxford: Polity Press.
- Holland, J. H. 1998. Emergence: from chaos to order. Ma.: Addison Wesley. ISBN 0-201-14943-5
- Holloway, T., Bozicevic, M., & Borner, K. 2007. Analyzing and visualizing the semantic coverage of Wikipedia and its Authors: Research Article Complexity, 12(3): 10.
- Jackson, M. C. 2000. Systems Approaches to Management. London: Kluwer Academic.
- Johnson, J. P. 2002. Open Source Software: Private Provision of a Public Good. Journal of Economics and Management Strategy, 11(4): 637-662.
- Jones, M. & Karsten, H. 2003. Review: Structuration Theory and Information Systems Research: 38. Cambridge: The Judge Institute of Management, Cambridge University.
- Kaptelinin, V. & Nardi, B. 1997. Activity Theory: Basic Concepts and Applications, CHI 97.
- Kauffman, S. 2000. Investigations. New York: Oxford. ISBN: 0-19-512104-x
- Kauffman, S. A. 1993. The Origins of Order: Self Organization and Selection in Evolution: Oxford University Press.
- Kauffman, S. A. 1996. At home in the Universe: The Search for Laws of Complexity. London: Penguin.
- Keeney, B. P. 1987. Aesthetics of Change: Guilford. ISBN: 1572308303
- Kittur, A., Suh, B., Pendleton, B. A., & Chi, E. H. 2007. He Says, She says: Conflict and coordination in Wikipedia. Paper presented at the Computer/Human Interaction 2007, San Jose USA.
- Lampert, A., Dale, R., & Paris, C. 2005. Classifying Speech Acts Using Verbal Response Modes. Sydney, Australia: CSIRO.
- Leont'ev, A. N. 1978. Activity, Consciousness and Personality. Engelwood Cliffs: Prentice Hall.
- Lewis, R. 1997. An Activity Theory framework to explore distributed communities. Journal of Computer Assisted Learning, 13: 8.
- Lih, A. 2004. Wikipedia as Participatory Journalism: Reliable Sources? Metrics for evaluating collaborative media as a news resource, 5th International Symposium on Online Journalism. University of Texas, Austin.
- Lin, Y. 2004. Hacking Practices and Software Development: A Social Worlds Analysis of ICT Innovation and the Role of Free/Libre Open Source Software. University of York, UK.
- Lizardo, O. 2004. The Cognitive Boundaries of Bourdieu's Habitus. Journal for the Theory of Social Behaviour, 34(4): 29.
- Lorenz, E. N. 2001. The Essence of Chaos (4 ed.). Seattle: University of Washington Press. ISBN: 0-295-97514-8
- Luhmann, N. 1990. Essays on Self Reference. New York: Columbia University Press. ISBN: 0231063687
- Luhmann, N. 1995. Social Systems. Stanford: Stanford University Press. ISBN: 0804726256
- Luthiger Stoll, B. 2005. Fun and Software Development. Paper presented at the First International Conference on Open Source Systems, Genova.
- Magnus, P. D. 2006. Epistemology and Wikipedia, North American Computing and Philosophy Conference. Troy, New York.
- McHenry, R.; The Faith based Encyclopaedia, <http://www.techcentralstation.com/111504A.html>
- Meissner, J. O. 2005. Relationship Quality in the Context of Computer Mediated Communication. Basel: University of Basel.
- Meyer, B.; Defence and Illustration of Wikipedia, [http://www.eiffel.com/general/monthly\\_column/2006/January.html](http://www.eiffel.com/general/monthly_column/2006/January.html)
- Michailakis, D. 1995. Review Essay : Law as an Autopoietic System. Acta Sociologica, 38: 14.
- Miettinen, R. 2006. Epistemology of Transformative Material Activity: John Dewey's Pragmatism and Cultural-Historical Activity Theory. Journal for the Theory of Social Behaviour, 36(4): 18.
- Mingers, J. 2002. Are Social Systems Autopoietic? Assessing Luhmann's Social Theory. Sociological review, 50(2).
- Mitchell, D.; Insider Editing at Wikipedia, <http://www.nytimes.com/2005/12/24/technology/24online.ready.html?ex=1293080400&en=431aff478b00239e&ei=5090&partner=rssuserland&emc=rss>

- Moody, G.; This time, it'll be a Wikipedia written by experts, <http://technology.guardian.co.uk/weekly/story/0,,1818630,00.html>
- Muffatto, M. & Faldani, M. 2003. Open Source as a Complex Adaptive System. Emergence: Complexity and Organization, 5(3), 1521-3250: 83-100.
- Muller, H. 1994. Luhmann's Systems Theory as a Theory of Modernity. New German Critique, 61(15): 39.
- Novack, T. P. & Hoffman, D. L. 1997. Measuring the Flow Experience Among Web Users: Vanderbilt University.
- Nupedia; Nupedia.com Editorial Policy Guidelines; 12 April, 2007, <http://web.archive.org/web/20010607080354/www.nupedia.com/policy.shtml>
- O'Reilly, T.; Open Source Paradigm Shift, [http://tim.oreilly.com/articles/paradigmshift\\_0504.html](http://tim.oreilly.com/articles/paradigmshift_0504.html)
- Oliver, R. I. & Bearden, W. O. 1985. Crossover Effects in the Theory of Reasoned Action: A Moderating influence attempt. The Journal of Consumer Research, 12(3), ISSN: 1531-8125: 324-340.
- Orlikowski, W. J. & Robey, D. 1991. Information Technology and the Structuring of Organizations. Information Systems Research, 2(2): 26.
- Osterloh, M., Rota, S., & Kuster, B. November, 2002. Open Source Software Production: Climbing on the Shoulders of Giants: 34: University of Zurich.
- Paulsen, N. & Hernes, T. (Eds.). 2003. Managing Boundaries in Organizations. NY: Palgrave Macmillan.
- Pedersen, S. T. June, 2002. Open Source and the Network Society 27: University of Aarhus, Denmark.
- Peterson, G. R. 2006. Species of Emergence. Zygon, 41(3): 22.
- Poole, M. S. & DeSanctis, G. 2002. Structuration Theory in Information Systems Research, Organization Science Winter Conference. Winter Park, CO.
- Porter, M. E. 1996. What is Strategy? Harvard Business Review(Nov- Dec).
- Postmes, T., Haslam, A., & Swaab, R. I. 2005. Social Influence in Small Groups: An interactive model of social identity formation. European Review of Social Psychology, 16, ISSN: 1046-3283: 1-42.
- Preece, J., Maloney-Krichmar, D., & Abras, C. 2003. History and emergence of online communities. In B. Wellman (Ed.), Encyclopaedia of Community: 11: Berkshire Publishing Group, Sage.
- Raymond, E. S.; Homesteading the Noosphere; 10, 3,
- Raymond, E. S. 2001. Cathedral and the Bazaar: Musings on Linux and Open Source by an Accidental Revolutionary (Revised ed.): O'Reilly Media. ISBN 0596001088
- Reagle, J.; A Case of Mutual Aid: Wikipedia, Politeness, and Perspective Taking, <http://reagle.org/joseph/2004/agree/wikip-agree.html>
- Ren, Y., Kraut, R., & Kiesler, S. 2007. Applying Common Identity and Bond Theory to Design of On-line Communities. Organization Studies, 28(3): 377-408.
- Ricci, A., Omicini, A., & Denti, E. n.d. Activity Theory as a framework for MAS coordination. Bologna: DEIS Universita di Bologna.
- Richardson, K. A. 2002a. Methodological Implications of a Complex Systems Approach to Sociality: Some further remarks. Journal of Artificial Societies and Social Simulation, 5(2), ISSN: 1460-7425.
- Richardson, K. A. 2002b. On the Limits of Bottom Up Computer Simulation: Towards a Non-linear Modeling Culture. Paper presented at the 36th Hawaii International Conference on Systems Science, Hawaii.
- Riehle, D. 2006. How and Why Wikipedia Works: An Interview with Angela Beesley, Elisabeth Bauer and Kizu Naoko, Wikisym 06. Odense, Denmark.
- Rossi, M. A. 2004. Decoding the "Free/Open Source (F/OSS) Puzzle" - a Survey of Theoretical and Empirical Contributions: 42. Sienna: University of Sienna.
- Rossi, M. A. April, 2004. Decoding the "Free/Open Source (F/OSS) Puzzle" - a Survey of Theoretical and Empirical Contributions: 42: University of Sienna.
- Rullani, F. 2005. The debate and the community: The reflexive identity concept and the FLOSS community case 32: Santa'Ann School of Advanced Studies.
- Sanger, L.; Britannica or Nupedia? The Future of Free Encyclopedias, <http://www.kuro5hin.org/story/2001/7/25/103136/121>
- Sanger, L.; Why Wikipedia must Jettison its Anti-elitism; Friday December 31 2004, <http://www.kuro5hin.org/story/2004/12/30/142458/25>
- Sanger, L.; The Early History of Nupedia and Wikipedia: A Memoir, file:///C:/data/surrey/Case%20Studies/wikipedia/early%20history%20of%20nupedia%20and%20wikipedia.htm
- Sanger, L.; Toward a New Compendium of Knowledge, [http://www.citizendium.org/essay\\_shorter.html](http://www.citizendium.org/essay_shorter.html)

- Sanger, L.; The Nupedia myth, <http://talkback.zdnet.com/5208-10535-0.html?forumID=1&threadID=22228&messageID=422549&start=-39>
- Sanger, L.; Why Make room for experts in web 2, <http://www.citizendium.org/roomforexperts.html>
- Sanger, L.; The Early History of Nupedia and Wikipedia: A Memoir, file:///C:/data/surrey/Case%20Studies/wikipedia/early%20history%20of%20nupedia%20and%20wikipedia.htm
- Sanger, L.; How to Think about Strong Collaboration among Professionals, [http://www.citizendium.org/collab\\_prof.html](http://www.citizendium.org/collab_prof.html)
- Sanger, L.; Why the Citizendium Will (probably) succeed, <http://www.citizendium.org/whyczwillsucceed.html>
- Sawyer, K. R. 2001. Emergence in Sociology: Contemporary Philosophy of Mind and Some Implications for Sociology Theory. *American Journal of Sociology*, 107(3), ISSN: 0002-9602: 551-585.
- Sawyer, K. R. 2003. Artificial Societies: Multi-agent Systems and the Micro-macro Link in Sociological Theory. *Sociological Methods & Research*, 31, ISSN: 0049-1241: 38.
- Sawyer, K. R. 2005. *Social Emergence: Societies as Complex Systems*. Cambridge, UK: Cambridge University Press. ISBN: 0521606373
- Searle, J. R. 1969. *Speech Act: An Essay in the Philosophy of Language*. Cambridge: Cambridge University Press.
- Sheppard, B. H., Hartwick, J., & Warshaw, P. R. 1988. The Theory of Reasoned Action: A Meta analysis of past research with recommendations for Modifications and Future Research. *The Journal of Consumer Research*, 15(3), ISSN: 1531-8125: 325-343.
- Shrader, W. E. 2005. *The Metaphysics of Ontological Emergence*. University of Notre Dame.
- Smith-Lovin, L. 2003. Self, Identity, and interaction in an ecology of Identities.
- Smith, J. R., Terry, D. J., & Hogg, M. A. 2007. Social Identity and the attitude-behaviour Relationship: Effects of anonymity and accountability. *European Journal of Social Psychology*, 37, ISSN: 0046-2772: 239-257.
- Stanford Encyclopaedia of Philosophy. 2006. Emergent Properties, *Stanford Encyclopaedia of Philosophy*.
- Stewart, I. 1990. *Does God Play Dice - The New Mathematics of Chaos*: Penguin. ISBN: 0631232516
- Stiles, W. B. 1992. *Describing Talk: A Taxonomy of Verbal Response Modes*: Sage. ISBN 0803944659
- Stiles, W. B., Lyall, L. M., Knight, D. P., Ickes, W., Waung, M., Lowry Hall, C., & Primeau, B. E. 1997. Gender Differences in Verbal Presumptuousness and Attentiveness. *Personality and Social Psychology Bulletin*, 7: 13.
- Stryker, S. & Burke, P. J. 2000. The Past, Present and Future of an Identity Theory. *Social Psychology Quarterly*, 63(4): 13.
- Stvilia, B., Twidale, M. B., Gasser, L., & Smith, L. C. 2005. Information Quality Discussions in Wikipedia. Illinois: Graduate School of Library and Information Science.
- Terry, D. J., Hogg, M. A., & White, K. M. 1999. The theory of planned behaviour: Self-identity, social identity and group norms. *British Journal of Social Psychology*, 38, ISSN 0144-6665: 225-244.
- Teubner, G. 1989. How the Law Thinks: Toward a Constructivist Epistemology of Law *Law & Society Review*, 23(5): 31.
- Therborn, G. 2002. Back to Norms! On the Scope and Dynamics of Norms and Normative Action. *Current Sociology*, 50(6), ISSN 0011-3921: 17.
- Viegas, F., Wattenberg, M., & Dave, K. 2004. *Studying Cooperation and Conflict between Authors with history flow Visualizations*. Paper presented at the CHI2004, the premier international conference for human-computer interaction, Vienna, Austria.
- Viegas, F., Wattenberg, M., Kriss, J., & van Ham, F. 2007. Talk Before you Type: Coordination in Wikipedia, 40th Hawaii International Conference on System Science. Hawaii.
- Voss, J. 2005. Measuring Wikipedia, *ISSI 2005*. Stockholm.
- Vygotsky, L. S. 1962. *Thought and Language*. Cambridge, Mass: MIT Press.
- Walther, J. B. & Burgoon, J. K. 1992. Relational Communication in Computer Mediated Interaction. *Human Communication Research*, 19(1): 38.
- Walther, J. B. 1995. Relational Aspects of Computer Mediated Communication: Experimental Observations Over Time. *Organization Science*, 6(2): 17.
- Watson, A. 2005. Reputation in Open Source: 30: College of Business Administration Northeastern University.

- Weik, E.; Working Relationships, A Meta view on Structure and Agency; 1, 7, <http://theoryandscience.icaap.org/content/vol7.1/weik.html>
- Widjaja, I. & Balbo, S. 2005. Structuration of Activity: A view on Human Activity, OZCHI 2005. Canberra, Australia.
- Wilkinson, D. M. & Huberman, B. A. 2007. Assessing the value of cooperation in wikipedia. Palo Alto: HP Labs.
- Williams, G. P. 1997. Chaos Theory Tamed. Washington D.C: Joseph Henry Press. ISBN: 0-309-06351-5
- Williamson, O. E. & Winter, S. G. 1993. The Nature of The Firm: Origins, Evolution and Development. New York: Oxford University Press.
- Zeitlyn, D. 2003. Gift economies in the development of open source software: anthropological reflections. Research Policy 32(7): 1287-1291