

# Archetypal Narratives in Social Machines: Approaching Sociality through Prosopography

Sécolène Tarte  
E-Research Centre  
University of Oxford  
Oxford, United Kingdom  
segolene.tarte@oerc.ox.ac.uk

Pip Willcox  
Bodleian Libraries  
University of Oxford  
Oxford, United Kingdom  
pip.willcox@bodleian.ox.ac.uk

Hugh Glaser  
Seme4  
& Ethos Valuable Outcomes  
London, United Kingdom  
hugh@glasers.org

David De Roure  
E-Research Centre  
University of Oxford  
Oxford, United Kingdom  
david.deroure@oerc.ox.ac.uk

## ABSTRACT

Introducing Social Machines as web-enabled entities integrating social energies and computational powers into a socio-technical system (whether purposeful or not) where social dynamics animate communities, this paper proposes a theoretical framework in which to observe them. Attempting to strike a balance between the roles of humans and non-humans, and aware of the difficulties that this heterogeneity presents, we propose to approach the questions of capturing the social dynamics of a social machine through prosopography. Prosopography is a method, used in particular by historians, that allows to systematically study a collection of biographies, be they of persons, artefacts, infrastructures of groups thereof. Systematization is achieved through designing an appropriate questionnaire to gather homogeneous data across the biographies. Our questionnaire design relies on the identification of five archetypal elements in biographical narratives. Illustrating our method with three examples, we demonstrate how our archetypal narratives have the potential to describe at least aspects of the social dynamics in social machines.

## Categories and Subject Descriptors

H.4.m [Information Systems Applications]: Miscellaneous

## General Terms

Theory, Human Factors, Design

## Keywords

Social Machine, socio-technical systems, narratives

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [Permissions@acm.org](mailto:Permissions@acm.org).

*WebSci '15*, June 28 - July 01 2015, Oxford, United Kingdom  
Copyright 2015 ACM ISBN 978-1-4503-3672-7/15/06 ...\$15.00  
DOI: <http://dx.doi.org/10.1145/2786451.2786471>

## 1. INTRODUCTION

The concept of social machine first emerged in 1999, when Prof. Sir T. Berners-Lee wrote: “Real life is and must be full of all kinds of social constraint—the very processes from which ‘society’ arises. Computers can help if we use them to create abstract *social machines* on the Web: processes in which people do the creative work and the machine does the administration. [...] The stage is set for an evolutionary growth of new social engines. The ability to create new forms of social processes would be given to the world at large, and development would be rapid [...]” [2, pp172–5]. Taking this quote as a starting point, the SOCIAM project<sup>1</sup> endeavours to observe, understand, and engineer such social machines. The social machines under study are all mediated through the web, whether they live on the web or are supported by web-enabled mobile devices. And, although it could be argued that any tool that enables or generates some kind of social interaction or organisation (e.g., a plough) would *sensu stricto* qualify as a social machine, our gaze is here firmly set on web-enabled socio-technical systems. Their observation, which is the focus of this paper, as well as their engineering are therefore pertinent to web sciences.

## 2. SOCIAL MACHINES: SCOPE

Let us first consider how the thinking about social machines has evolved with regard to devising a definition. A first attempt was to define social machines through characterising them [14]. The resulting classification, although valuable as a tool, is very much biased towards the design and engineering of the mechanisms that can be found in social machines and pays only little attention to the social dynamics that inhabit and move them. Concomitantly, some efforts have been made to encompass the dynamic dimension of social machines by understanding them as living entities that define ecosystems, have life cycles, and typically exhibit some kind of emergent properties [4]. As ecosystems, they can be nested, as well as overlapping and interacting (whether collaborating or competing). Further, all the elements in a social machine –people, data, algorithms, digital objects, and infrastructure– are participants in the social

<sup>1</sup>Full project title: Theory and Practice of Social Machines; website: [sociam.org](http://sociam.org)

machine, rather than users (which would imply consumption rather than action) or processors (which would imply the sole providing of a predefined service). The major conceptual shift operated through social machines is thus that they democratically allow all involved, animate or inanimate<sup>2</sup>, to have a participatory role. This does not exclude participants from having the possibility to choose to solely consume, or to solely provide a service, simply it allows them all to actively partake.

It is therefore helpful at this stage to reflect upon how we have been communicating about social machines, and how our use of traditional terminologies might be implicitly biasing us towards the machines in social machines, thereby unwittingly diverting our attention away from the social dimension of social machines. Indeed, in our discourse about specific social machines, we often resort to calling them, as a shortcut, by a name designating the results they produce or the main platform that supports them. Examples include Wikipedia, Zooniverse, and Twitter; and although these are shortcuts to designate social machines that are all rather different in nature, designating them so unintentionally occults the social dynamics that feed them and keeps us focused on the main platforms that support them, to the exclusion of all the other platforms that might be used routinely but not quite as visibly (see section 5.1 for such an example). This observation might seem a petty terminological issue, but terminology reinforces our natural cognitive biases, and, in this very case, it keeps us asking traditional engineering questions such as: “who are the users?” “why do people participate?” “what are the boundaries of a social machine?” Those questions are perfectly valid ones to ask of centralised service-providing systems or even if the intent of the system is to “harness the power of the crowd”, but if we rather consider that **social machines are entities integrating social energies and computational powers into a socio-technical system (whether purposeful or not) where social dynamics animate communities**, and if we adopt names that identify these social communities to designate the social machines, then those questions often become inapt. Indeed, communities are not defined by their boundaries or by the natural boundaries of the tools they use –in fact a mix and match of a number of tools is usually likely, technology being appropriated by individuals as well as by communities– nor do communities have users; and questions related to people’s partaking, although remaining valid, take on a much more interactive and social colouring.

It is on the basis of this view that **social machines are entities that are socio-centric at least as much as they are techno-centric** that our methodology of observation of social machines was developed. In an attempt to capture the dynamic aspects of social machines characterised by the circulation and creation of ideas, knowledge, and values typical of sociality, it was previously proposed to consider the narratives that unfold within and about social machines [15]. Building upon this work, we contend that by specifying the type of narratives to gather in order to study and understand the social dynamics of social machines, it will be possible to systematize data collection to obtain homogeneous data across the heterogeneous entities of a social machine. The specific narratives that are required are bi-

<sup>2</sup>We use the terms with their etymological meaning of respectively having breath, life (animate), or not (inanimate).

ographies; by proposing to harvest the biographies of the various types of entities within a social machine –biography of a unitary entity such as a person, a piece of data, an algorithm, a digital object, a piece of infrastructure, or biography of a composite entity such as a group of people, data, algorithms, digital objects, infrastructure, be they homogeneous groups or heterogeneous groups– we are actually setting the grounds for a prosopographical approach to the observation of sociality in social machines.

### 3. BIOGRAPHIES & PROSOPOGRAPHY

Biography is a familiar, perhaps instinctive, narrative form. It results from our pattern-making impulse which makes stories of series of occurrences between people and things, sets them in a wider context, and describes “the ways meanings and values are accumulated and transformed” [8, p172].

Despite its etymology<sup>3</sup> a biography is not restricted to studies of human lives, but has come to mean a “themed narrative history of a specific subject in any of various written, recorded, or visual media”<sup>4</sup>. A biographical approach has been usefully applied to artefacts as diverse as a Fijian whales’ teeth necklace [8, p171] and e-Infrastructure [12]. As a biography of a person might include a description of the family and the culture into which she was born, the biography of a thing incorporates its technological predecessors and its design alongside its “life-history” (Tringham’s concept [16] –cited in [8, p169]) and fall from use. Following the recommendations of Pollock & Williams’ Biographies of Artefacts framework [12], we therefore adopt the notion that the design of a piece of software, a buttons in a software, an app, infrastructure, or an abstracted computational entity are an intrinsic part of their biographies. Biographies cannot encompass an entire life history. They are of necessity partial — temporally, spatially, and in point of view.

From its earliest extant formalization in the works of Cornelius Nepos<sup>5</sup> and Plutarch<sup>6</sup>, the biographical form has been used comparatively, setting narratives of great people next to one another to instruct readers in, for example, cultural relativity and ethics. As the number of biographies and the complexity of entities they describe grows, as in our work on social machines, more structure than simple proximity is needed to make sense of the whole. Accordingly we have adopted a prosopographical method, using its framework to illuminate the process of meaning emerging from social actions [8, p170] through collecting and collating homogeneous data on the heterogeneous biographies of social machine entities.

A prosopography can be a collection of life studies, or a collective biography<sup>7</sup>. Most pertinently to our work, it “in-

<sup>3</sup>βίο- life + -γραφία writing

<sup>4</sup>“biography, n.”. OED Online. March 2015. Oxford University Press: <http://www.oed.com/view/Entry/19219>

<sup>5</sup>c. 110 BC – c. 25 BC; Roman biographer of, amongst others, Cato the Censor

<sup>6</sup>c. AD 46 – AD 120; Greek author of *Parallel Lives* - Βίοι Παράλληλοι, a set of pairs of biographies that aim to put forth qualities and failings of famous men; a number of the 23 pairs of surviving biographies from the *Parallel Lives* can be consulted at [http://www.perseus.tufts.edu/cache/perscoll\\_Greco-Roman.html](http://www.perseus.tufts.edu/cache/perscoll_Greco-Roman.html)

<sup>7</sup>“prosopography, n.”. OED Online. March 2015. Oxford University Press: <http://www.oed.com/view/Entry/153010>

tegrates more or less large numbers of descriptive individual biographical studies into quantitative and statistic research” [17]. Starting “from concrete data” a prosopography “aims at understanding general phenomena”[17]: it is not an aim or end in itself, but a method for collating data at scale.

The prosopographical method requires the formulation of a questionnaire to systematize the collection of data. The questionnaire will have a specific research objective or objectives. Answering it for each biography creates a single model from our heterogeneous data, conjuring a comprehensible uniformity out of the data.

When applied to a collection of biographies of people the prosopographical method might, for example, establish place of birth, education undertaken, and place, age and cause of death. Our questionnaire takes the design principle of archetypal narratives, borrowing from Jungian archetypal premises [10] and from Propp’s morphology of Russian folktales [13].

Our aim in using this method is to find a way to describe and so to recognize and ultimately create social machines. While interesting as a study, if it is not useful it has failed. By nature, then, the development of our prosopographical method is iterative, and our selection and description of these archetypal narratives tested against the results they provide.

We use archetypal narratives to characterize and to decompose the biographical narratives of and in social machines. Through their lens, we can view non-mutually exclusive combinations and concatenations of the archetypal narratives that form a biography, and account for the dynamic interactions between archetypes within their mutable forms. Archetypes are not building blocks in primary colours, but rather denote proclivities or persuasions at a particular point in time or space, or through one narrative point of view.

#### 4. ARCHETYPAL NARRATIVES

The concept of archetypal narratives has been explored in varied ways. One way was Propp’s, who, with a structuralist approach, looked for the typical characters and plot lines in Russian folktales [13]. His work allows to identify building block in folktales. Eco, later adopted a similar approach to analyze Fleming’s Bond books [5]. The aim these works achieve is to find typical structures in the narratives they study. Following a different line of investigation, that of the human psyche, Carl Jung also developed a theory of archetypes. His archetypes however derive from the notion of collective unconsciousness [10] and aim to understand human thought and behaviour. Jung’s archetypes never apply directly to characterize people, instead they combine, shift, morph, and interact, giving some perspective onto why a given person thinks and behaves how they do. The aim is not classificatory, but a tool for observation and diagnosis. The fact that these archetypes emanate from the collective unconscious and have the ability to mutate, combine, compete and meld together bestows all its power onto Jung’s method, relating the person to their environment and to the dynamics within and around them.

Our search for archetypal elements of biographical narratives in the context of a prosopographical observation of social machines is inspired by Propp & Eco and Jung’s work. We are looking to identify elements of narrative that are typical of biographies, but also that have the ability to mutate,

combine, compete and meld together, thereby capturing dynamic elements, and revealing connections beyond the unitary level. This presents the advantage of allowing us to explore community dynamics, going beyond individual incentives, which are often the focus of socially-motivated observation in Web Sciences (e.g. [1, 6])

So that in each of our archetypal narratives, the main character of the biography is the entity whose life story in the social machine is being told. (i.e., person, piece of infrastructure, data, algorithm, digital object, as well as any-size groups thereof or systems encompassing a range of these units); we shall call this entity *the protagonist* from now on. With a small focus group composed of a physicist, a historian, a digital humanist, an engineer, and a social scientist all involved with social machines in some capacity, we first drew up a list of the kinds of elements of biographical narratives we could identified as repeatedly occurring in social machines; we then whittled that list down to five archetypal elements of narratives in the biographies of social machines. These were then put to the test by asking SOCIAM project members to each think of the biography of an entity in a social machine, and to assess whether our archetypes helped to qualify those biographies. This is of course an iterative process, so we will continually be assessing our archetypes against specific cases. For now, the list of five archetypal elements of narratives in the biographies of social machines seems to have stabilised as follows:

**{LAND}** A canvas, or settings, or landscape [11] - that forms the baseline on which the biographies develop. This is the humus or the substrate on which all other archetypes can express themselves. Examples of what can constitute part of such a landscape include:

- Routine, such as a repetitive task
- The mundane, such as gossip and banter
- The constraints & restrictions that apply in all forms, and particularly those arising from technology
- The traditions & conventions that apply in all forms, and particularly those arising from social interactions

**{OUT\_txf}** A change or transformation of the protagonist’s environment. The environment is made of all the entities external to the protagonist. So for an inanimate protagonist, all animate participants in the social machine that might (or not) be interacting with it are part of the environment. Examples of changes of the environment include:

- Evolution or repurposing, such as: a forum being changed into a chat room (for a person as a protagonist); or a forum member becoming a moderator (for a digital object as a protagonist)
- Power shifts, such as those provoked by the commercialization of a platform (for a person as a protagonist), or those provoked by changes in a legal system (e.g., for a digital object as a protagonist, the change of law affecting people’s rights to access it)

**{IN\_txf}** A change or transformation of the protagonist’s internal state. Examples of changes of the internal state include:

- Metamorphosis and change, such as a shift in belief system for a person as a protagonist, or the renaming of a discussion functionality from forum to chat room for the discussion space functionality as a protagonist
- Discovery and learning, such as those experienced by many people participating in citizen science projects

**{I-O-INTERACT\_txf}** A change or transformation of the interaction between the protagonist’s internal state and its environment; these are often paired with **{OUT\_txf}** or with **{IN\_txf}**. Examples of changes of the interaction between internal state and environment include:

- Adaptation, such as those where the protagonist operates a shift in their mode of interaction with their environment due to environmental changes (so here **{I-O-INTERACT\_txf}** is paired with **{OUT\_txf}**)
- Diversion (including subversion), where implicit affordances are made explicit, and thus change the nature of the interaction, such as a system used for a purpose not initially intended, from the perspective of any protagonist interacting with the system in line with its original explicit intent

**{EVENT}** An event that involves the protagonist. Examples of events include:

- Encounters
- Beginnings & endings

These archetypes might seem to overlap or to not be conceivable as stand-alone archetypal narratives –this is actually a very good sign, as it is in the ways they combine that the descriptive powers of this approach lies. These archetypal narratives are also mutable by nature, of course, and therefore we have further identified the following qualities, or textural characteristics, that every one of the five archetypal narratives can display, and that can change over time (i.e., they can have a proleptic/prospective value –as evaluated in the moment– and a number of retrospective values –as evaluated after the fact).

**<coin>** A degree of coincidence; the degree to which things happen at the same time seemingly by chance;  
in the range: ordinary  $\longleftrightarrow$  random

**<plan>** A degree of planned-ness; the degree to which things have been pre-scripted;  
in the range: unplanned  $\longleftrightarrow$  planned

**<gen>** A degree of generative-ness; this relates to kairotic time, i.e., to the weight, significance, importance given to something through the way it is perceived in time and the causal relation it holds with subsequent happenings, the potential it holds for further development;  
in the range: static  $\longleftrightarrow$  dynamic

**<span>** A timespan;  
in the range: punctual  $\longleftrightarrow$  durative

Coincidence and planned-ness have been separated to allow for design-enabled coincidences, which will have a high level of coincidence (**<coin>**) and of planned-ness (**<plan>**);

serendipity however will occur with a high degree of coincidence (**<coin>**) and a low degree of planned-ness (**<<plan>**); intention will be characterised by a high degree of planned-ness (**<plan>**).

For example, an **{EVENT}** will naturally tend to be more dynamic than static (**<gen>**) and more punctual than durative (**<<span>**), and a **{LAND}** will naturally tend to be more static than dynamic (**<<gen>**) and more durative than punctual (**<span>**).

## 5. CASE STUDIES

To illustrate how these archetypes map onto biographies of participants in social machines, and because the biographies of inanimate participants might not be quite as self-explanatory as those of people, the examples below outline such biographies with their corresponding archetypes and how they combine. To facilitate the structuring of these biographies, we shall outline them as a list of “Moments in a lifetime” [7] (rather than as purely discursive narratives); the archetypes however might express themselves at such a Moment, but also across several Moments or as ways to transition from one Moment to another.

### 5.1 Biography of a collection of images

In this example, the protagonist is a collection of images of galaxies now known as pea galaxies, or for short “the green peas”, through the social machine of the classifiers in Galaxy Zoo. The Moments are chronological but only a sample of the actual lifetime –by nature; they are followed by a list of the archetypes they express, and a graphical representation of these expressions can be found in Fig. 5.1.

**Moment (0).** In the context of the Galaxy Zoo project<sup>8</sup> (a citizen science project currently in its fourth version, originally launched in its first version in 2007), a large number of images from the Sloan Digital Sky Survey (SDSS<sup>9</sup>) were put online in order for participants in the Galaxy Zoo project to classify the galaxies in the images. At this stage, our protagonist, the collection of images of green peas, does not yet exist as such. The images of this collection have not been noticed as part of a collection yet, but we can already describe the kind of environment that the collection will live in. It is the social machine that is made of the following elements: the whole Galaxy Zoo platform with its elements of interface that allow Galaxy Zoo classifiers, aka Zooites, to access and classify the images as well as the forum where Zooites have the opportunity to discuss and exchange amongst themselves and with the moderators and scientists; the Zooites along with the moderators –the Zookeepers– and the scientists; the SDSS website, where all the images come from and to which the Zooites regularly refer in conversations; all the images and collections of images; and the on-going repetitive tasks of classification, punctuated by remarks and exchange on images, which can range from short aesthetic interjections to long exchanges that are rich in scientific content. **{LAND}**

<sup>8</sup>Website: <http://www.galaxyzoo.org/>, one of the many projects that form the Zooniverse (<https://www.zooniverse.org/>).

<sup>9</sup>Website: <http://skyserver.sdss.org/>

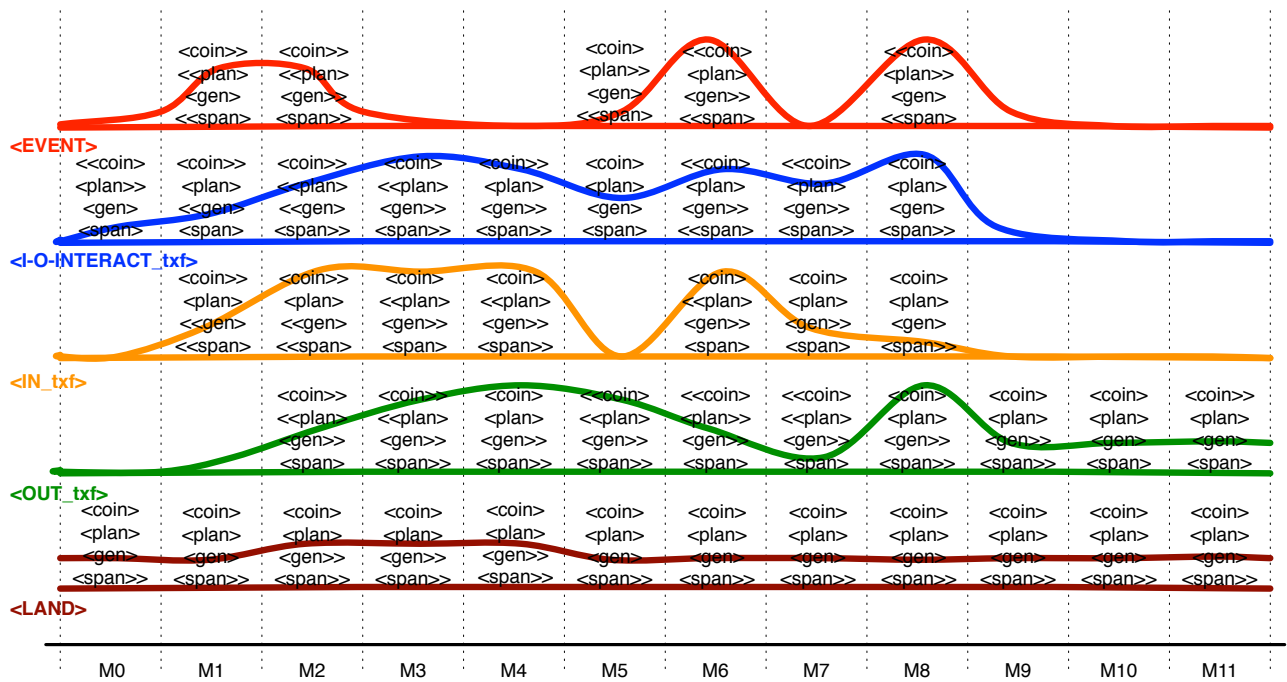


Figure 1: The Moments in the life of the collection of images of Green Peas in the Zooites’ social machine, and their corresponding archetypes: if the life of the collection is compared to a 5-strand thread, where each strand represents an archetypal element of a biographical narrative, then this graph outlines the thickness of each strand, for each of the Moments in the life-history of the collection. Each expression of each archetype is further qualified through its four textural characteristics.

*Moment (1).* On 26th July 2007, in the “stunning sights” section of the forum a Zooite starts a “green galaxy”<sup>10</sup> thread, which (retrospectively) seems to be the first that flags an image of what will later be known as green peas.  
 $\{LAND\}, \{I-O-INTERACT\_txf\}, \{IN\_txf\}, \{EVENT\}$

*Moment (2).* On 12th August 2007, a Zooite starts a thread on the forum: “Give peas a chance!”<sup>11</sup>, where Zooites are invited to collect images of what looks like green peas. Before that date, a number of posts in the “weird and wonderful” section of the forum regularly flagged green-looking galaxies - the posters of these are invited by Zookeeper Alice to contribute their green finds to the “Give peas a chance!” thread. The collection is born and named. Zooites actively contribute images of green peas to the collection, which from then on grows very fast. Around it, Zooites exchange puns and jokes, as well as more serious observations about the galaxies in the images they’re collecting. The Zooites collecting green peas name themselves “the Peas Corps” (hereafter the PC-Zooites).  
 $\{LAND\}, \{I-O-INTERACT\_txf\}, \{IN\_txf\}, \{EVENT\}, \{OUT\_txf\}$

*Moment (3).* Amidst many more jokes and pea-based puns, PC-Zooites actively discuss the criteria for inclusion in the collection, scientists and Zookeepers chip in. PC-Zooites develop programs that search for green peas through the SDSS, in order to compile lists of images of green peas, and

post them in the Give Peas a chance forum. The collection, started as a casual harvest on the Galaxy Zoo, now also grows out of data harvested directly on SDSS.  
 $\{LAND\}, \{I-O-INTERACT\_txf\}, \{IN\_txf\}, \{OUT\_txf\}$

*Moment (4).* The growing collection has sparked further interest in the scientists monitoring the Galaxy Zoo, and on 8th July 2008, a new thread called “Peas project”<sup>12</sup> was started with the collection of green peas at its centre, moderated by Yale student astronomer Carie who explicitly aims to collect “these interesting galaxies”, and to publish the findings around this new type of galaxy. Specific criteria of recruitment for the inclusion of images of green-looking objects in the collection are formulated. (PC-)Zooites keep contributing images to the collection, some are included in the collection, some are not when they don’t meet the criteria.  
 $\{LAND\}, \{I-O-INTERACT\_txf\}, \{IN\_txf\}, \{OUT\_txf\}$

*Moment (5).* On 7th July 2009, the collection of green peas is the subject of a blog post<sup>13</sup>; this post details the hunt for the green peas and how the PC-Zooites deployed computational methods to search the SDSS.  
 $\{LAND\}, \{I-O-INTERACT\_txf\}, \{EVENT\}, \{OUT\_txf\}$

*Moment (6).* The collection of green peas is the subject of an academic paper [3], published in November 2009. The collection of green peas has given rise to the discovery of a

<sup>10</sup><http://www.galaxyzooforum.org/index.php?topic=158.0>

<sup>11</sup><http://www.galaxyzooforum.org/index.php?topic=3638.0>

<sup>12</sup>Forum thread: <http://tinyurl.com/lk6mcu7>

<sup>13</sup>Blog post: <http://wp.me/p2mbJY-fX>

type of galaxy now called “pea galaxy”. The actual collection of green peas now gives official way to the concept of a pea galaxy.

{LAND},{I-O-INTERACT\_tx},{IN\_tx},{EVENT},{OUT\_tx}

**Moment (7).** On 10th December 2009, a Wikipedia “pea galaxy” entry is created by a PC-Zooite.

{LAND},{I-O-INTERACT\_tx},{IN\_tx},{OUT\_tx}

**Moment (8).** On 30th September 2012 the “Peas project” thread is archived- all this time, Zooites have been contributing images of green peas. The natural home of the collection is now the SDSS database, where the images originally came from, and where “pea galaxies” can be retrieved thanks to the criteria that define them.

{LAND},{I-O-INTERACT\_tx},{IN\_tx},{EVENT},{OUT\_tx}

**Moment (9).** The pea galaxy, and therefore the Galaxy Zoo green peas collection, has obviously been giving ideas to scientists, who have now discovered “green beans”<sup>14</sup>.

{LAND},{OUT\_tx}

**Moment (10).** On 2nd March 2014, a sample from the collection of green peas features as “Amazing Galaxy of the Week” on the Daily Zooniverse<sup>15</sup>.

{LAND},{OUT\_tx}

**Moment (11).** In March 2015, the authors of the current paper write a biography of the collection of green peas, visiting the forum archives, visioning some of the images, in pursuit of the understanding of how this narrative within and about the Zooites’ social machine is a telling clue to its social dynamics.

{LAND},{OUT\_tx}

## 5.2 Biography of a URI

The protagonists in this example are Linked Data URIs. Linked Data URIs are used by Semantic Web technologies to identify concepts, and adhere to the following Linked Data principles<sup>16</sup>:

1. Use URIs as names for things
2. Use HTTP URIs so that people can look up those names.
3. When someone looks up a URI, provide useful information, using the standards (RDF\*, SPARQL)
4. Include links to other URIs, so that they can discover more things.

Principles 1-3 are the important ones here, and so we consider what look like normal HTTP URLs that identify concepts by being names for them and return a document that has “useful” information in the context of the Semantic Web.

We describe the changes an URI can go through as a set of Moments. During each of the Moments, the URI interacts with one or more agents which are often part of a Social

Machine such as a group of people and software tools deciding to agree that the URI identifies the same concept as another URI. One such example would be the context of the Pelagios project [9], where scholars contribute place names and geographical locations, along with the Ancient evidence for those, in order to map them in the Ancient World. The Moments below are general descriptions of Moments, some of which are followed by more specific examples. They are further followed by the list of archetypes they express; as they are all assumed to occur within a context such as the Pelagios project, they will all be assumed to express the {LAND} archetype, specific to their context. They are not in a strict order, and can occur multiple times, but there are some temporal dependencies - for example Moment 0 usually has to happen before other Moments. Fig. 5.2 represents the expressions of the archetypes for each Moment.

**Moment (0).** The URI is just a gleam in the eye - some person or automatic agent coins a Concept that might be communicated.

[Romulus and Remus decide to build Rome]

{LAND},{OUT\_tx}

**Moment (1).** The URI is constructed - in our case, a Linked Data URI, which is a string of the appropriate form.

[DBpedia (the Linked Data site related to Wikipedia) is planned and the URI scheme constructs <http://dbpedia.org/resource/Rome> (hereafter dbpedia:Rome) as the URI]

{LAND},{IN\_tx},{EVENT}

**Moment (2).** A document describing the Concept is created.

[The DBpedia extraction software processes Wikipedia, and creates the RDF as required]

{LAND},{OUT\_tx}

**Moment (3).** The document is placed at the location of the URI.

[DBpedia is built, and so dbpedia:Rome now resolves to an RDF (and HTML) document]

{LAND}{I-O-INTERACT\_tx},{IN\_tx}

**Moment (4).** The URI is used in a other documents.

[[http://dbpedia.org/resource/Province\\_of\\_Rome](http://dbpedia.org/resource/Province_of_Rome) refers to the dbpedia:Rome]

{LAND},{I-O-INTERACT\_tx},{OUT\_tx}

**Moment (5).** A person or agent accesses the document using the URI.

[The author of this paper accesses dbpedia:Rome to verify it is correct]

{LAND},{EVENT}

**Moment (6).** A person or agent decides that the URI is the “sameAs” another URI.

[Sam decides that

[http://dbpedia.org/resource/Province\\_of\\_Rome](http://dbpedia.org/resource/Province_of_Rome) is the “sameAs” dbpedia:Rome]

[Pat decides that <http://sws.geonames.org/3169071/> is the “sameAs” dbpedia:Rome]

{LAND},{I-O-INTERACT\_tx},{IN\_tx},{EVENT},{OUT\_tx}

<sup>14</sup><http://www.sci-news.com/astromy/article00763.html>

<sup>15</sup><http://daily.zooniverse.org/2014/05/02/green-pea-galaxies/>

<sup>16</sup><http://www.w3.org/DesignIssues/LinkedData.html>

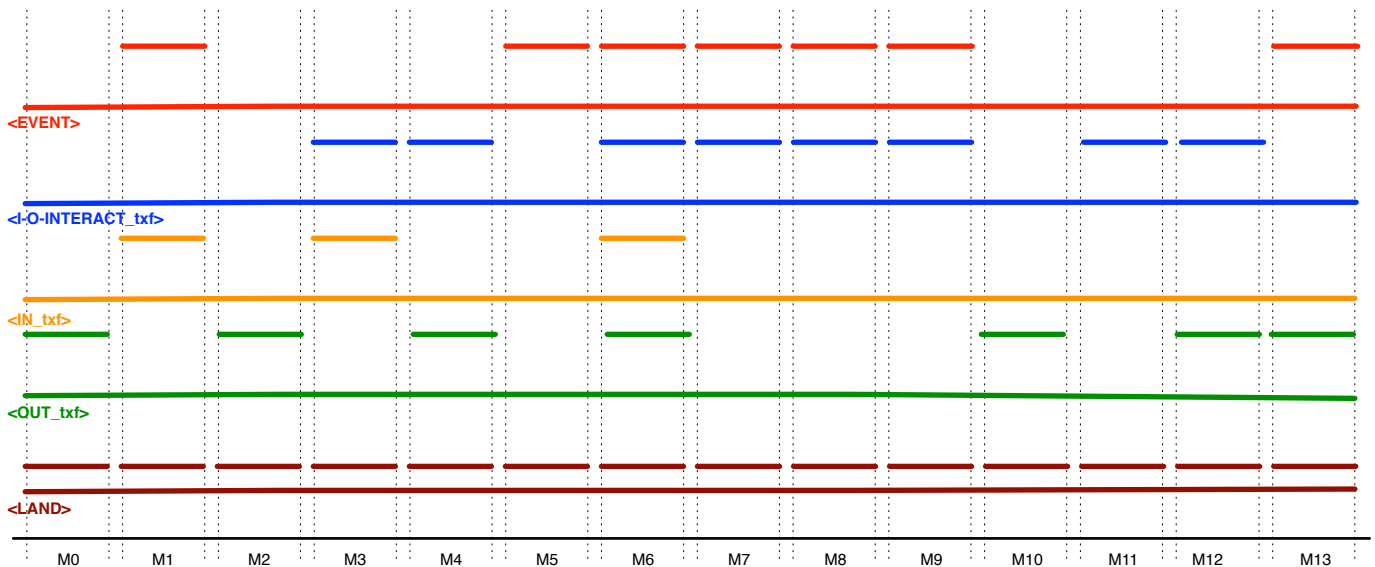


Figure 2: Typical Moments in the life of a URI within a social machine, and their corresponding archetypes. In this representation, the textural characteristics are not specified, because the Moments outlined here are generic and every lived Moment will express its archetypes with its own specific textural characteristics.

*Moment (7).* A person or agent asserts that the URI is the “sameAs” another URI in a document or system, such as sameAs.org.  
 [Sam creates a document that says that: dbpedia:Rome owl:sameAs http://dbpedia.org/resource/Province\_of\_Rome]  
 {LAND},{I-O-INTERACT\_tx},{EVENT}

*Moment (8).* A person or agent decides that the URI is “differentFrom” another URI.  
 [Les decided that http://sws.geonames.org/3169071/ is “differentFrom” dbpedia:Rome (the geopolitical entity of Rome as capital is different from the municipality)- note that this assertion conflicts with the M6 decisions]  
 {LAND},{I-O-INTERACT\_tx},{EVENT}

*Moment (9).* A person or agent asserts that the URI is “differentFrom” another URI in a document or system, such as differentFrom.org.  
 [Les publishes a document with the “differentFrom” data from M8 in, although differentFrom.org refuses to accept it]  
 {LAND},{I-O-INTERACT\_tx},{EVENT}

*Moment (10).* The document describing the URI is deleted.  
 {LAND},{OUT\_tx}

*Moment (11).* The URI is looked up and no document is found.  
 {LAND},{I-O-INTERACT\_tx}

*Moment (12).* All the documents that had the URI in it get deleted.  
 {LAND},{I-O-INTERACT\_tx},{OUT\_tx}

*Moment (13).* Every person and agent has “forgotten” about the Concept.  
 [The world ends]  
 {LAND},{OUT\_tx},{EVENT}

### 5.3 Biography of the retweet action

In this example, the principle of in-Twitter citation of tweets is the protagonist, as we consider retweets (RT). Interestingly, this story also illustrates an example of the humans programming the social machine, through the use of RT (retweet), MT (modified tweet), and HT (hat tip). Further, this story gives a social account, through the retweet action, of the evolution of a piece of infrastructure (the Twitter platform) used in a large number of social machines. Fig. 5.3 represents the expressions of the archetypes for each Moment.

*Moment (0).* A Twitter “echo” is imagined. At 11:28 on 7 March 2007, Narendra Rocherolle (@narendra), wanting to encourage attendance at an SXSW event<sup>17</sup>, wrote “please twitter ‘echo’ this msg even if you can’t attend”<sup>18</sup>. This appears to have been the invention of the concept of re-(micro)blogging on Twitter<sup>19</sup>.  
 {LAND}

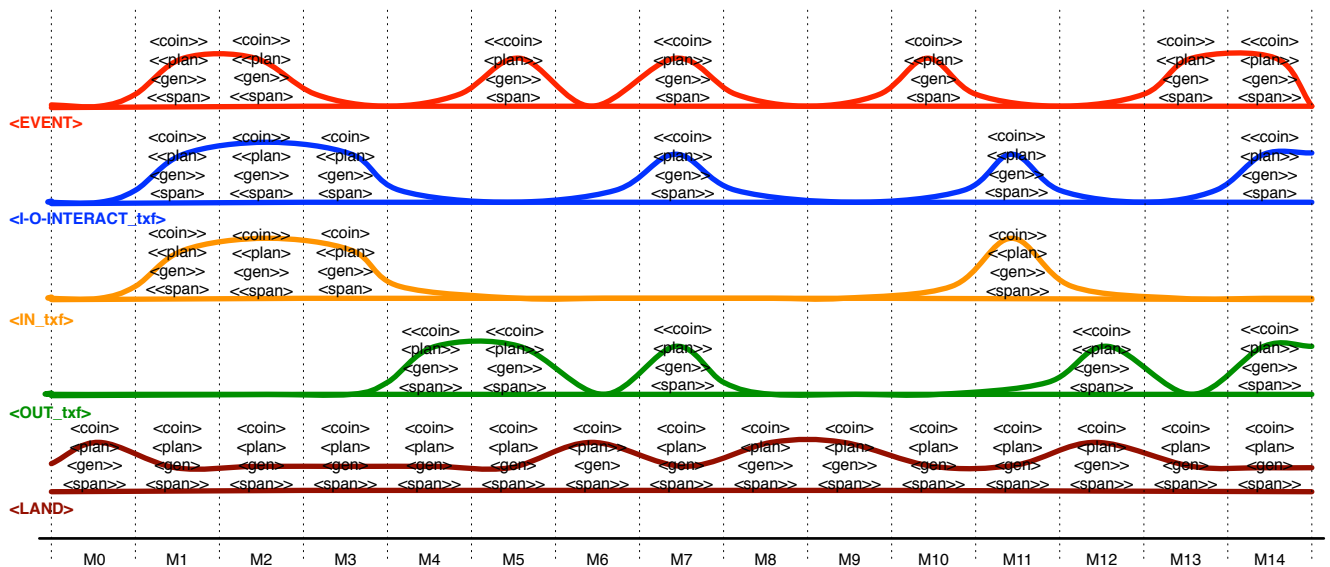
*Moment (1).* The term “retweet” is coined. At 20:33 on 17 Apr 2007, Eric Rice (@ericrice) coined this usage of the word “ReTweet”<sup>20</sup> when quoting a reply he had received from another microblogger, Jesse Malthus (@jmalthus) about Web 2.0 and social media. The term “retweet” had been used previously, but apparently only to refer to reposting one’s

<sup>17</sup><http://tinyurl.com/puh9pko>

<sup>18</sup><https://twitter.com/narendra/status/5911334>

<sup>19</sup><http://en.wikipedia.org/wiki/Reblogging>

<sup>20</sup><https://twitter.com/ericrice/status/31669791>



**Figure 3: The Moments in the life of in-Twitter citations of tweets in the Tweeters' social machine, and their corresponding archetypes. Each expression of each archetype is further qualified through its four textural characteristics.**

own tweets<sup>21</sup>.

{LAND}, {IN\_txf}, {I-O-INTERACT\_txf}, {EVENT}

**Moment (2).** The abbreviation “RT” is coined. At 12:15 on 25 January 2008, @TDavid made the first recorded use of “RT” as an abbreviated form of “retweet”<sup>22</sup>, followed by the @<username> formation<sup>23</sup> when retweeting news from @BreakingNewsOn about a fire in Las Vegas.

{LAND}, {IN\_txf}, {I-O-INTERACT\_txf}, {EVENT}

**Moment (3).** Retweets are adopted more widely by the Twitter community. During the 2009 Iranian presidential election<sup>24</sup>, the use of retweeting increased as microbloggers retweeted reports and observations sent via SMS<sup>25</sup>.

{LAND}, {IN\_txf}, {I-O-INTERACT\_txf}

**Moment (4).** A Twitter project to develop a retweet button is announced. At 20:54 on 13 August 2009, co-founder and Creative Director of Twitter Biz Stone publishes an article on Twitter about the development of a retweet button. The blogpost recognizes the invention and adoption of retweeting by the Twitter community: “Retweeting is a great example of Twitter teaching us what it wants to be”<sup>26</sup>. Stone notes that the “RT @<username>” form is “a bit cumbersome”, and outlines the method in development to allow retweeted tweets to appear directly in the timelines of retweeters and their followers. This saves precious characters of the 140 allowed by Twitter. It also creates an easy metric for inferring

<sup>21</sup>Seward, Zachary (15 October 2013). “The first-ever hashtag, @-reply and retweet, as Twitter users invented them”: <http://tinyurl.com/kaotoj6>. Quartz

<sup>22</sup><http://tinyurl.com/n8d9ygg>

<sup>23</sup><https://twitter.com/TDavid/status/641334922>

<sup>24</sup>[http://en.wikipedia.org/wiki/President\\_of\\_Iran](http://en.wikipedia.org/wiki/President_of_Iran)

<sup>25</sup><http://en.wikipedia.org/wiki/Reblogging>

<sup>26</sup>Stone, Biz (13 August 2009). “Project Retweet: Phase One”. Twitter blog: <http://tinyurl.com/mavdynz>

popularity or notoriety.

{LAND}, {OUT\_txf}

**Moment (5).** The retweet button is launched. At 00:50 on 6 November 2009, Biz Stone announced that a Retweet button had been activated on “a very small number of accounts”<sup>27</sup> to make “forwarding a particularly interesting tweet to all your followers very easy”. The retweet button’s reception was mixed<sup>28</sup>, with support for its ease of use and complaints that people were seeing retweets they did not want by people whose timelines they follow.

{LAND}, {OUT\_txf}, {EVENT}

**Moment (6).** The newly launched retweet button has detractors. At 12.49 on 10 November 2009, Jacob Mullins (@jacob) tweets: “Don’t think I like twitter Beta ‘Retweet’ feature - can’t comment on the tweet and I like the RT at the beginning, qualifies it from begin”<sup>29</sup>.

{LAND}

**Moment (7).** The retweet mute function is rolled out. At 11:37 on 10 November 2009, co-founded of Twitter Ev Williams (@ev) tweets that a “per-user” mute function is available, to avoid seeing unwanted retweets in timelines<sup>30</sup>. The ability for a tweeter to silence a nuisance (to her mind) retweeter is a blunt instrument, blocking the medium and not the message: blocking one messenger’s path does not preclude tweets from a particular account or on a particular theme from arriving through another person’s retweets.

{LAND}, {OUT\_txf}, {I-O-INTERACT\_txf}, {EVENT}

<sup>27</sup>Stone, Biz (13 August 2009). “Retweet Limited Rollout”, Twitter blog: <http://tinyurl.com/17e83rg>

<sup>28</sup>Siegler, M.G. (10 November 2009). “Hate It Or Love It, Twitter’s New Retweet Style Is Rolling Out”. TechCrunch: <http://tinyurl.com/39b4ww1>

<sup>29</sup><https://twitter.com/jacob/status/5598900212>

<sup>30</sup><https://twitter.com/ev/status/5597147435>



**Moment (8).** The newly launched retweet button has supporters. At 12:55 on 10 November 2009, Matt Galligan (@mg) tweets: “Yup, the official Retweet user experience is pretty solid. Liking it so far.”<sup>31</sup>.  
{LAND}

**Moment (9).** The retweet button versus “RT” wars continue. On 31 May 2011, Nathan Bransford blogs the question “RT or the Retweet Button?”, weighing the merits of each, and concluding “RT @Username Tweet is so 2010”<sup>32</sup>. The tone is light, and the title interrogative; yet even the use of “should” reflects the seriousness with which retweets are debated.  
{LAND}

**Moment (10).** A retweet record is set. At 20:16 on 6 November 2012, as US presidential election results were being announced<sup>33</sup>, Barack Obama tweeted a photograph of himself hugging his wife Michelle, with the words “Four more years”<sup>34</sup>. Within 24 hours it had been retweeted over half a million times<sup>35</sup>. Its retweets currently stand at 752,471. His next tweet said, “This happened because of you. Thank you.”<sup>36</sup>  
{LAND},{EVENT}

**Moment (11).** Meta-retweeting as tool for analysis. On 12 November 2012, Peter Bray blogs about the life span of a tweet, noting that without access logs a way to judge when a tweet has been “consumed” (Twitter’s term) is through its retweets. A retweet is a transaction: “I’ve consumed this tweet and find it valuable enough to pass along to others”<sup>37</sup>. Bray notes the spread of similar functions to Facebook and other networks.  
{LAND},{IN\_txf},{I-O-INTERACT\_txf}

**Moment (12).** There is rearguard action in support of “traditional” retweets. On 3 September 2013, Ray Beckerman describes how the retweet button, inserting tweets into the timelines of a tweeter’s followers, takes the “social” out of “social media”: “[a] eliminating conversation and interaction, [b] insisting on blind rubber stamping, [c] preventing you from letting your friends know you’ve honored them, [d] preventing you from knowing your friends have honored you, [e] making you invisible, [f] making it harder for you to meet new friends with similar interests, and [g] removing any clear indication of your identity to your existing friends.”<sup>38</sup>  
{LAND},{OUT\_txf}

**Moment (13).** The retweeted photograph record is broken. At 19:06 on 2 March 2014, television host Ellen DeGeneres tweets a group selfie taken by Bradley Cooper of actors at the Oscars ceremony which she had hosted<sup>39</sup>. Within an hour it was retweeted over a million times, and over two million by the end of the ceremony, causing Twitter to crash<sup>40</sup>. Its retweets currently stand at 3,362,388. She later commented, “We really just made history. We’re all winners tonight.”  
{LAND},{EVENT}

**Moment (14).** “Retweet” becomes “Share” on iOS. At 14:48 on 27 March 2014, the UK’s Daily Mail newspaper reports that Twitter has replaced its “iconic” retweet button with a “Facebook-style” “Share with followers” option, to the distress of some tweeters<sup>41</sup>.  
{LAND},{OUT\_txf},{I-O-INTERACT\_txf},{EVENT}

In all the case studies above, the inanimate protagonist undergoes transformations and experiences events that affect it and the whole social machine it lives in. In the case of the retweet action, the main transformation that occurs yields a subtle re-coding through the emergence of a new convention, that is now considered etiquette; with the “green peas” the transformation yields a new scientific discovery, and knowledge enrichment; in the case of URLs, it is wider and easier access to information that is enacted. In all these examples, the archetypes express themselves with reference to the protagonist (whether it’s a unitary entity or a composite one). When participants meet, their biographies share Moments, and for each of these Moments archetypes will be expressed through the biographies of the protagonists involved, that is with the various protagonists’ points of view (so the archetypal composition of a shared Moment will differ depending on which protagonist’s point of view is assumed). These meetings of biographies, these mergings, and interminglings are what adds another dynamic social dimension to the already dynamic life-histories of the protagonists; they can turn into collective biographies and then allow observation across scales.

## 6. CONCLUSION

This paper proposes a theoretical framework for the observation of web-enabled social machines, and in particular their social dimension, which is one of their main source of energy; this framework is informed both by the social sciences (Biographies of Artefacts framework), and the tried and tested methodological approach to understanding relationships between entities in History, prosopography. Setting out to design a way to interrogate social machines and their biographies, we have identified five archetypal elements of biographical narratives that allow us to uniformly interrogate heterogeneous entities across scales (entities can be units as well as composites). The power of these archetypes as a descriptive tool is inherent in their ability to combine and influence each other, allowing for internal dynamics (in a

<sup>31</sup><https://twitter.com/mg/status/5599038967>

<sup>32</sup>Bradford, Nathan (31 May 2011). “Should You Use RT or the Retweet Button on Twitter?”. <http://tinyurl.com/m9wtj43>. Nathan Bransford author

<sup>33</sup>US presidentials 2012 on wikipedia: <http://tinyurl.com/d19s5g>

<sup>34</sup>B. Obama twitter status: <http://tinyurl.com/c2q44tw>

<sup>35</sup>News report on T3: <http://tinyurl.com/acrh245>

<sup>36</sup>B. Obama twitter status: <http://tinyurl.com/k3w9oxu>

<sup>37</sup>Bray, Peter (12 November 2012). “When Is My Tweet’s Prime of Life? (A brief statistical interlude.)”: <http://tinyurl.com/ljye9mx>. The Moz Blog

<sup>38</sup>Beckerman, Ray (3 September 2013). Ray’s 2.0: <http://tinyurl.com/osedjh3>

<sup>39</sup>The Ellen Show status: <http://tinyurl.com/19w9fppd>

<sup>40</sup>Guardian article: <http://tinyurl.com/17k8vmf>

<sup>41</sup>Woollaston, Victoria 27 March 2014. “End of the retweet? Twitter rumoured to be replacing its iconic button with a Facebook-style ‘share’ option.” <http://tinyurl.com/of23y63>. Daily Mail

way similar to Jung’s archetypes, which do not characterise people but rather the dynamics that govern, motivate, and animate them). The archetypal narratives therefore offer a much more flexible framework for analysis than any taxonomical framework. It would be tempting, based on these archetypal biographical narratives to derive a list of archetypal social machines. Caution should apply however, as it is in the plurality of the archetypal narratives and in the mutual influences they exert upon each other when they co-exist that the richness of this descriptive approach resides. The examples we provided demonstrate the kind of information that can be harvested within this framework.

Future work will have to tailor the harvesting techniques to the entities whose biographies are being gathered. So that, in practice, the next big question is: how to identify and then collect the data that will serve as evidence of the expression of one or another of the archetypal narratives in a given social machine? This is actually a non-trivial sampling problem, a sampling of big & wide data across heterogeneous entities, and it will require thoughtful and adaptive instrumentation. Finally, beyond observation, this framework, once the difficult instrumentation questions have been tackled, has the potential to also supply valuable information for the design of social machines.

## 7. NOTE

All web addresses herein: last visited on 20th March 2015.

## 8. ACKNOWLEDGMENTS

This work is supported under SOCIAM: The Theory and Practice of Social Machines, a programme funded by the UK Engineering and Physical Sciences Research Council (EPSRC) under grant number EP/J017728/1, and a collaboration between the Universities of Edinburgh, Oxford, and Southampton.

The authors further wish to thank the following people for the fruitful discussions that helped the ideas in this paper to grow: Dr Robert Simpson, Dr Victoria Van Hynning, Dr David Murray-Rust, Dr Mark Hartswood, Prof. Stuart Anderson, Dr Grant Miller.

## 9. REFERENCES

- [1] L. M. Aiello, R. Schifanella, and B. State. Reading the source code of social ties. In F. Menczer, J. Hendler, and W. Dutton, editors, *WebSci '14- Proceedings of the 2014 ACM conference on Web science*, pages 139–48. ACM, 2014.
- [2] T. Berners-Lee and M. Fischetti. *Weaving the web: The original design and ultimate destiny of the world wide web by its inventor*. Harper, San Francisco, 1999.
- [3] C. Cardamone, K. Schawinski, M. Sarzi, S. P. Bamford, N. Bennert, C. M. Urry, C. Lintott, W. C. Keel, J. Parejko, R. C. Nichol, D. Thomas, D. Andreescu, P. Murray, M. J. Raddick, A. Slosar, A. Szalay, and J. Vandenberg. Galaxy Zoo Green Peas: discovery of a class of compact extremely star-forming galaxies. *Monthly Notices of the Royal Astronomical Society*, 399:1191–1205, Nov. 2009.
- [4] D. De Roure, C. Hooper, M. Meredith-Lobay, K. Page, S. Tarte, D. Cruickshank, and C. De Roure. Observing social machines part 1: what to observe? In *Proceedings of the 22nd international conference on World Wide Web companion (WWW '13 Companion), 1st International Web Observatory Workshop (WOW2013)*, 2013.
- [5] U. Eco. The narrative structure in Fleming. In T. B. Bernard Waitesr and G. Martin, editors, *Popular Culture: Past and Present*, pages 242–263. Routledge, 1982.
- [6] P. L. Gay, S. Brown, A. D. Huang, C. Lehan, and Moon Zoo Team. Behaviors and Motivations observed in the Zooniverse. *AGU Fall Meeting Abstracts*, page A581, Dec. 2010.
- [7] H. Glaser. On the parallel evaluation of expressions in functional languages. In E. Chiricozzi and A. d’Amico, editors, *Parallel processing and Applications*, pages 327–332. Elsevier Science Publishers, September 1987. Address: L’Aquila, Italy.
- [8] C. Gosden and Y. Marshall. The cultural biography of objects. *World Archaeology*, 31(2):pp. 169–178, 1999.
- [9] L. Isaksen, R. Simon, E. T. Barker, and P. de Soto Cañamares. Pelagios and the emerging graph of ancient world data. In F. Menczer, J. Hendler, and W. Dutton, editors, *WebSci '14- Proceedings of the 2014 ACM conference on Web science*, pages 197–201. ACM, 2014.
- [10] C. G. Jung. *The Archetypes and the Collective Unconscious (Collected Works of C. G. Jung)*. Routledge, 2 edition, 1991.
- [11] D. Murray-Rust, S. Tarte, M. Hartswood, and O. Green. On wayfaring in social machines. In *WWW '15: Companion Publication of the 24th International World Wide Web Conference*. ACM, (forthcoming 2015).
- [12] N. Pollock and R. Williams. E-infrastructures: How do we know and understand them: Strategic ethnography and the biography of artefacts. *Computer Supported Cooperative Work (CSCW)*, 19(6):521–556, 2010.
- [13] V. Propp. *Morphology of the folktale*. University of Texas Press, Austin, 1968.
- [14] N. R. Shadbolt, D. A. Smith, E. Simperl, M. Van Kleek, Y. Yang, and W. Hall. Towards a classification framework for social machines. In *SOCM2013: The Theory and Practice of Social Machines*, Rio de Janeiro, Brazil, 2013.
- [15] S. M. Tarte, D. De Roure, and P. Willcox. Working out the plot: the role of stories in social machines. In E. Simperl, N. Contractor, J. Hendler, and N. Shadbolt, editors, *WWW '14: Companion Publication of the 23rd International World Wide Web Conference*, pages 909–914. IW3C2, ACM, 2014.
- [16] R. Tringham. Archaeological houses, households, housework and the home. In D. N. Benjamin, D. Stea, and D. Saile, editors, *The Home: Words, Interpretations, Meanings, and Environments*, pages 79–107. Aldershot, Avebury, 1995.
- [17] K. Verboven, M. Carlier, and J. Dumolyn. A short manual to the art of prosopography. In K. S. B. Keats-Rohan, editor, *Prosopography Approaches and Applications. A Handbook*, volume 13 of *Prosopographica et Genealogica*. Occasional Publications UPR, Oxford, 2007.