Cohere: Towards Web 2.0 Argumentation

Simon BUCKINGHAM SHUM
Knowledge Media Institute, The Open University, UK

Abstract: Students, researchers and professional analysts lack effective tools to make personal and collective sense of problems while working in distributed teams. Central to this work is the process of sharing—and contesting—interpretations via different forms of argument. How does the “Web 2.0” paradigm challenge us to deliver useful, usable tools for online argumentation? This paper reviews the current state of the art in Web Argumentation, describes key features of the Web 2.0 orientation, and identifies some of the tensions that must be negotiated in bringing these worlds together. It then describes how these design principles are interpreted in Cohere, a web tool for social bookmarking, idea-linking, and argument visualization.

Keywords: argumentation tools; argument visualization; usability; Web 2.0

1. Introduction: The Need for Distributed, Collective Sensemaking Tools

The context in which we find ourselves presents problems on a global scale which will require negotiation and collaboration across national, cultural and intellectual boundaries. This, I suggest, presents both major challenges and unique opportunities for us, as the community dedicated to understanding computational support for argumentation: there is a particular need to provide tools with languages for communities to track how they agree and disagree in principled ways.

While previous work has focused on real time dialogue and argument mapping in contexts such as e-science teams [1] and personnel rescue [2], this paper focuses specifically on the challenge of designing engaging, powerful tools for distributed, primarily asynchronous, argumentation. The paper begins by reflecting on the emergence of what has been dubbed the “Web 2.0” paradigm, identifying the kinds of expectations that Web users increasingly bring. We then survey the current state of the art in Web Argumentation tools, concluding that to date, only one attempt has been made to develop a tool that displays the qualities that have made social information sharing and structuring tools so successful. We then describe how Web 2.0 principles, as introduced, have informed the design of a tool called Cohere, concluding with a vision of how COMMA researchers might extend or interoperate with it.
2. The Web 2.0 Paradigm

A lot is being written about the Web 2.0 paradigm, a term first dubbed in 2004 [3]. While some dismiss it as marketing hype, it does serve as a useful umbrella term to cover significant new patterns of behaviour on the Web. There are many lists of the key characteristics of Web 2.0, not all of which are relevant to our concerns (e.g. e-business models). In this section we select several characteristics for their impact on the user experience of collective information structuring. Together these present a challenge to the design of practical Web Argumentation tools, given the expectations that users now have from their everyday experience of the Web. If we cannot create tools within the new landscape, argumentation tools will remain a much smaller niche than they should be — and as this paper seeks to demonstrate, need be.

2.1. Simple but Engaging Multimedia User Interfaces

The World Wide Web has established itself as the default platform for delivering interactive information systems to professionals and the public. Although early Web applications lacked the elegance and interactivity of desktop applications due to the need for the client to communicate every state change to the server, the gap is closing rapidly with the emergence of good graphic design principles, controlled layout and stylesheet management, and critically, so-called Rich Internet Applications: interactive multimedia capabilities such as Adobe Flash embedded as standard browser plugins, and approaches such as AJAX (Asynchronous JavaScript And XML) for caching local data to increase the responsiveness of the user interface [4]. Users increasingly expect Web applications to have a clean, uncluttered look, and to be as responsive as offline tools. Given a choice of Web offerings, the user experience can determine whether or not a tool is adopted.

2.2. Emergent, Not Predefined, Structure and Semantics

Argumentation focuses on a particular kind of semantic structure for organising elements. Of central interest, therefore, is the Web 2.0 emphasis away from predefined information organizing schemes, towards self-organised, community indexing (‘tagging’) of elements, resulting in so-called “folksonomies” that can be rendered as tag clouds and other visualizations. Persuading ‘normal people’ (in contrast to information scientists or ontology engineers) to create structured, sometimes high quality, metadata was previously thought impossible, and the success and limits of this approach is now the subject of a new research field that studies collaborative tagging patterns, e.g. [5].

Another way in which this emphasis expresses itself is in the new generation of tools that make it easy to publish one’s opinion of the world. Free, remotely hosted blogging tools such as Wordpress and Blogger make it very easy for non-technical users to create a personally tailored journal or diary and syndicate their ideas. Blogs demonstrate one way to negotiate the formality gulf successfully, providing expressive freedom (essentially, traditional prose and graphics), with just enough structure to reap some benefits of hypertext (entries are addressable as URLs, timestamped, tagged, and syndicated as web feeds – see below). The limitation of blogging at present is that like the Web at large, there are no semantics on the links between sites, thus failing to
provide any support for an analyst who wants to gain some kind of overview of the relationships between posts.

2.3. Social Networks

Web 2.0 applications are dominated, although not exclusively restricted to, sites that explicitly seek to connect people with people, via the artifacts that they share. They are designed such that the greater the numbers of people who participate, the higher the return on effort invested. Social tools provide a range of ways in which users are made aware of peer activity, for instance, alerting when another user ‘touches’ your material (e.g. by reusing it, making it a favourite, tagging it), or by mining social network structure to suggest contacts in a professional network. Social tools also provide mechanisms for building reputation, from the trivial (how many “friends” one has), to more meaningful indices, such as authority based on the quality of material or feedback that a user posts, or professional endorsements.

2.4. Data Interoperability, Mashups and Embedded Content

A core idea behind the Web 2.0 paradigm is access to data over the web from multiple applications. Web feeds using RSS and Atom have become the lingua franca for publishing and subscribing to XML data in a simple manner that many non-technical users now handle daily. Public APIs and web services enable the more sophisticated access that enterprise architectures require, while semantic web services promise to overlay ontologies on these layers so that they can be configured according to function. “Mashups” of data sources fuse disparate datasets around common elements (e.g. geolocation, person, date, product), often accessed via customisable user interfaces such as Google Maps [6]. Mashups may need to be crafted by a programmer, but others can be generated by end-users, given a sufficiently flexible environment. The results of a search may bring together data in new ways.

The phenomenal growth of web applications such as Google Maps, YouTube, Flickr and Slideshare is in part due to the ease with which users can embed remotely hosted material in their own websites. By providing users with the ‘snippet’ code (which may be HTML or JavaScript), such applications empower users to in turn provide their readers with attractively presented access to the material, which can in turn be embedded by those readers in their sites. The material thus spreads ‘virally’, as the links to a resource increase: it is no longer necessary to visit a web page to access its content.

3. Web Argumentation Tools

A significant strand in COMMA research focuses on the design, implementation and evaluation of practical software tools for creating and analysing arguments. Following the entity-relationship modelling paradigm that lends itself so well to software, as well as the work of pioneering argument and evidence mapping theorists such as Wigmore and Toulmin, these tools provide a way to construct arguments as structures comprising semantically linked elements taken from one or more argumentation schemes. The argument structures may be left implicit behind text-centric user interfaces, or rendered explicitly as trees or networks to help the author and reader visualize and edit the
The intended users of such tools include members of the public engaged in a public consultations and societal debate [7], students or educators in a learning context [8], lawyers [9], and analysts in many other walks of professional life such as public policy [10] and scholarly publishing [11]. Research in this field examines issues including the translation of argumentation theory into computable representations [12], the nature of expert fluency with such tools [13, 14], and empirical studies of the tools’ usage in all of the above domains.

In light of the high design standards and new possibilities that the Web 2.0 paradigm sets, it is clear that existing tools have limitations. First, there are desktop applications like Compendium [30] and Rationale [15] with high quality user interfaces refined through the feedback from their extensive user communities: however, these are limited to publishing read-only maps to the Web, either as JPEG images, or as interactive image maps. Single user applications like CmapTools which have been migrated to ‘groupware’ versions provide remote editing of maps, but do not exploit the Web 2.0 functions described above.

Finally and most relevant, there are a number of Web-native applications, designed from the start to support large scale, multi-user construction. Some websites now provide a very simple structure for structuring the two sides of a debate, while others provide a more articulated argumentation language. Beginning with the least structured, we see the emergence of sites such as Debatepedia, which is modelled on Wikipedia, providing a debating resource showing unstructured prose arguments for and against a particular proposal, demarcated in two columns [16]. CoPe_it! [17] is designed for community deliberation, and provides a way to synchronise views between IBIS graphs (it also integrates with Compendium in this respect), an IBIS outline tree, and a conventional threaded discussion forum. CoPe_it! also provides a mechanism to evaluate the strength of a position, and so represents another interesting development. Its interaction design is at present rather rudimentary compared to Web 2.0 interfaces. It does not have an end-user customisable semantics, interoperability with existing Web data sources, or mechanisms to syndicate content outside the application.

Parmenides is designed to support web-based policy consultation with the public, and incorporates a formal model of argumentation [18]. It provides a forms-based, questionnaire interface to elicit views from the user, populating an argumentation structure, which it then reasons over to elicit further views. Parmenides enforces a particular argument ontology (it was not designed as a social web application) and does not appear to support any other Web 2.0 characteristics.

ClaiMaker [19] was a Web 1.0 era application, developed in our own prior work modelling the claims and arguments in research literatures. ClaiMaker, and its sister tool ClaimSpotter [20], provided vehicles for us to validate empirically the usability of the data model and a number of user interface paradigms. This has led us to carry the core data model through into Cohere, while relaxing the constraint that restricted users to the predefined classifications of nodes and links. Cohere’s visualizations are also versions of those first prototyped in ClaiMaker.

TruthMapping goes much further than this, aiming specifically at tackling some of the limitations of threaded discussion forums, with a clear distinction between unsupported premises, which when supported become claims, and a way to post rebuttals and responses to each of these [21]. DebateMapper uses a combined graphical and outline structure to map debates using the IBIS scheme, with contributions tagged as issues, positions and arguments [22]. None of these tools provides mechanisms to reason over the debate, or compute the strength of evidence: they serve to provide a
structured, collaboration medium that serves as a collective memory. DebateMapper in particular illustrates how most of the Web 2.0 principles above can be brought together, apart from open semantics, and is the closest work to what is reported here.

The ArgDF system [23] is the first web argumentation tool to adopt a Semantic Web architecture based around the W3C standard Resource Description Framework (RDF) for distributed data modelling and interchange. Moreover, ArgDF is probably the first interactive tool to ground its argument representation in the recently proposed Argument Interchange Format (AIF) [24]. This combination of AIF+RDF is a notable advance. However, while proving the conceptual and technical feasibility of a semantic web orientation for argumentation, it does not yet have a user community, and it cannot be regarded as a Web 2.0 application as defined above.

4. Towards Web 2.0 Argumentation: The Cohere system

We now describe how we are trying to incorporate the Web 2.0 principles introduced above to create an environment called Cohere [http://cohereweb.net] which aims to be semantically and technically open, provide an engaging user experience and social network, but provide enough structure to support argument analysis and visualization.

4.1. Emergent Semantics: Negotiating the Formalization Gulf

In any user-driven content website, the challenge is to keep entry barriers as low as possible to promote the growth of the community, yet maintain coherence of navigation and search, through the careful design of the data model and user interface. The combination of data model and user interface must seek the right balance between constraint and freedom. This Web 2.0 orientation might seem to be in tension with an environment designed to promote rigorous thinking and argumentation. Our approach is to start with relaxed constraints in order to foster engagement with the idea of structuring ideas in general, but provide tools to incrementally add structure as the user recognises the value that it adds in a given context.

Cohere is, therefore, styled to invite playful testing by people who may not first and foremost be interested in argumentation. Instead, the website invites them to make connections between ideas. This broader framing aims to meet the need of many sensemaking communities to express how ideas or resources are related (whether or not this is argumentative) in a way that goes beyond plain text blog postings, wikis or discussion forums. A typical pair of connected Ideas in Cohere is illustrated in Figure 1.

![Figure 1: Example of a user-defined connection between two Ideas in the Cohere system](image-url)

In Cohere, users are free to enter any text as an Idea and its detailed description. The examples seeding the database convey implicitly that Idea labels are generally short and succinct. Users are encouraged by the user interface to reuse existing Ideas, with an autocomplete menu dropping down as they type to show matching Ideas.
already published: as far as possible, we want them to describe the same Idea using the same label.

Users **must**, however, constrain their contributions by:
- creating labelled connections between Ideas (e.g. is an example of)
- reusing, or creating, a connection from a list of either positive or negative connections

Users can optionally:
- assign roles to Ideas (e.g. Scenario; Problem)
- add descriptive details (displayed when the Info icon is clicked)
- assign websites to Ideas (listed when the Idea is selected)

The Cohere data model is inherited from the ClaiMaker prototype [11]. The essence is summarised informally in Figure 2.

The provision of mechanisms to enable flexible linking of web resources around what we are calling Ideas is a goal shared by the Topic Maps community [25], whose data model is very close to that of Cohere. Intriguingly, the two were developed entirely independently, yet arrived at the same core data model, which we take to be a form of empirical validation. In the more formally defined, and more wide-ranging Topic Map Data Model, **topics** (=Cohere Ideas) point to one or more **resources** (=websites); topics can be linked by **associations** (=connections), and topics may play one or more **roles** within a given association (=roles). A Web 2.0 application called Fuzzy [26] is based on the Topic Map standard and shares some similarities with Cohere, as does the HyperTopic system [27]; neither, however, provide support for argumentation.

While not mandating that the user engage in argumentation, the language of argument is nonetheless at the heart of the system in terms of the default roles that Ideas can play in a connection (Figure 3), and the connection types offered to users (Figure 4).
Figure 3: Expanding an Idea to show the default, and user-defined, roles that the Idea can play in a connection.

Figure 4: Expanding the connection menu to show the default, and user-defined, connection types, which are classified as either supportive or challenging.
These default connection types are also leveraged in the predefined visualization filters offered by the Connection Net tab, described later (Figure 8). As these figures illustrate, however, users may choose to ignore the defaults and create their own connection language, and the roles Ideas can play. All connections are, nonetheless, classed as broadly positive or negative, which provides a way to express not only disagreement in the world of discourse, but could signify inhibitory influence (e.g. in biological or ecological systems modelling), or antagonistic relationships (e.g. in social networks). It is entirely up to the individual or team to agree on their modelling scheme.

4.2. Visualizing IBIS-Based Dialogue Mapping in Cohere

It can be seen from Figure 3 and Figure 4 that the default roles that an Idea can play in a connection are Questions, Answers, Pros and Cons. These are derived from the Issue-Based Information System (IBIS) developed by Rittel 28, and implemented in the Compendium tool referred to earlier. This is used to model what Walton and Krabbe [29] classified as deliberation dialogues over the pros and cons of possible courses of action to address a dilemma.

Our previous work has demonstrated the value of real time IBIS dialogue mapping in meetings, and the use of IBIS as an organising scheme around which an analyst can map, asynchronously, the structure of public policy debates which can then be published as read-only maps on the Web [30]. Cohere now provides a platform for collaborative deliberation and debate mapping over the internet, with primarily asynchronous use in mind to start with. (Real time mapping requires a tool like Compendium which has a user interface optimised for rapid mapping. However, it is our intention to optimise for real time mapping in the longer term, perhaps by adapting Compendium as an applet for Cohere).

4.3. Visualizing Argumentation Schemes and Critical Questions in Cohere

In related work [31], we have demonstrated how Walton’s argumentation schemes and associated Critical Questions, rendered as XML files in the Argument Markup Language [32], can be transformed into Compendium XML and expressed as IBIS structures in Compendium. The resulting argumentation scheme templates can now be modelled in Cohere as illustrated in Figure 5.

4.4. Social Networking and Reputation

All Ideas other than one’s own have their owner clearly indicated iconically. Clicking this displays the user profile, making it possible to learn more about the person behind the ideas. We are beginning to add metrics to make users aware when they arrive at the site how many active users there are, and what the most recently posted, reused and linked Ideas are. Web feeds in the future will enable users to request notification whenever one of their Ideas is embedded in someone else’s connection, or in someone else’s website (see below).
4.5. Interoperability: Web Data as Platform

Central to O’Reilly’s notion of Web 2.0 is the notion of web data as the platform on which many applications can compute. Cohere exposes and consumes data in a variety of ways:

- Publishing and importing XML Web feeds
- Importing XML data from the Compendium offline dialogue and argument mapping tool
- Embedding pointers to its data in other applications as URLs and HTML ‘snippets’
- Exposing data in a variety of standards to engage different communities
**Web feeds:** Cohere seeks to build on the significant effort that many users already invest in social bookmarking with folksonomic tagging tools such as del.icio.us, or in blogging with tools such as Blogger or Wordpress. These are currently two of the most dominant ways in which users share their views, and Cohere aims to leverage this by enabling users to import/refresh the Web feed (RSS or Atom) for any bookmarking or blogging site. Entries are converted into Ideas and annotated with the relevant URL, ready for linking. We are in the process of implementing an RSS feed so that users can track new Ideas as they are published. We plan to develop this capability, so that individual Ideas can be subscribed to, with alerts everytime someone connects to or from them.

**Ideas and views as URLs:** It is increasingly hard to find an artifact or building these days without a URL on it. The web depends on the URL as a way for non-technical users to connect web pages, save searches, and disseminate sites of interest via standard tools such as email, slides and wordprocessors. The design of URLs goes beyond cool top level domain names, to the art of URLs that communicate their content to people, in contrast to machine-generated addresses that have no obvious pattern.

It was considered essential, therefore, to make Cohere’s content addressable and accessible as URLs. This required the creation of a guest login status for non-registered users to successfully reach an address, and the design of a URL syntax that specified the visualization type and search criteria. The URL for an Idea, a triple, or a Connection List/Net is accessed by the user in what has become the conventional manner, by clicking on a URL icon to copy and paste the address that pops up.

**Embedding ideas and views in other websites:** Once a URL addressing scheme is in place, it becomes possible to provide such embeddable snippets for users, as introduced above. Pasting this `<iframe>` code into a web page creates an embedded, interactive view onto the Cohere database, which reproduces the buttons to get the URL and snippet code, to encourage further dissemination (Figure 6).

![Figure 6](image)

**Multiple Import/Export data formats:** By the time of the conference, we will have implemented further data formats for importing and exporting Cohere structures. A priority is to provide Argument Interchange Format compatibility, with other candidates being Topic Maps, Conceptual Graphs, and OWL.

4.6. **Mashup Visualizations**

Our objective is to help forge links not only between Ideas, but between the people publishing them. As Cohere starts to be used, it is inevitable that popular Ideas will be
duplicated: if the site is successful, we can expect many people to be working on the Idea Global Warming, or making reference to everyday concepts such as Capitalism or World Wide Web. We have started to design views that help render the structures that will result from many users working on common Ideas. This is a long term challenge, but Figure 7 shows the first tool called Connection Net, which uses a self-organising graph layout algorithm that can render all of one’s personal data, or filtered views of the world’s data. In particular, Ideas with a border are used by more than one person, and as shown, right-clicking on it enables the user to view all the owners of that Idea. In this way, just as folksonomies enable disparate users to discover related resources and people, Cohere aims to reveal new connections and users working on the same Idea, or perhaps more interestingly, making similar or contrasting connections.

![Image](image1.png)

Figure 7: The Connection Net view merges all matching Ideas in a single node, and lays out the graph automatically.

Filter buttons in the Connection Net view make use of the connection types, as shown in Figure 8. A number of saved filters are shown, for example, Contrast searches the database from a focal Idea on a specific subset of connection types of a contrasting nature, e.g. challenges, has counterexample, is inconsistent with, refutes. Users can define their own custom searches, and in the future will be able to save them as shown in the example buttons.

![Image](image2.png)

Figure 8: Semantic filter buttons that show only a subset of connection types from a focal Idea. The example shown is a Contrast search: rolling over it displays the connection semantics it will search on. User defined searches are issued from the Connection Search button on the left.
4.7. Implementation

Cohere is a ‘LAMP’ web application: Linux operating system, Apache HTTP server, MySQL database and scripted in PHP. The user interface exploits the AJAX approach to caching data in the browser to create a highly responsive interface, with few delays between action and feedback. Cascading Style Sheets are used extensively to control presentation. In addition, a Java applet from the Prefuse information visualization classes [33] has been customised to provide self-organising, interactive graph visualizations under the Connection Net tab. Compendium (op cit) serves as an offline mapping tool (a cross-platform Java desktop application with Apache Derby or MySQL database). Data is uploaded to Cohere currently using the Compendium XML scheme; see discussion on data interoperability for the other import/export formats that are being added. Maps in IBIS are translated into Cohere versions, website links are converted to hyperlinked Ideas, and labelled links become semantic connections.

5. Present Limitations, and Future Work

This paper has introduced the design rationale behind the introduction of Web 2.0 social software principles to Web Argumentation tools, in particular, when designing sensemaking support tools for distributed communities. At the time of writing there are about 50 alpha testers working in e-learning and professional contexts, and the tool is about to be launched. By the time of the conference, we will have initial quantitative and qualitative usage data to report, enabling us to reflect on the extent to which we have succeeded in creating an engaging tool for structuring and visualizing information and arguments, publishing Ideas, and discovering new intellectual peers.

Future work will support different argument layouts, the saving and loading of template libraries (currently managed via the offline Compendium tool), group permissions, and user evaluations in different domains. We are particularly keen to collaborate with professionals engaged in collective problem solving online.

We have begun to consider how argument modelling approaches and reasoning engines developed by other COMMA research groups could be integrated with Cohere. The obvious Web 2.0 way would be via simple Web XML feeds, but we can also envisage an appropriate service-oriented argumentation architecture [cf. 34] and/or data exchange via RDF+AIF [23], such that Cohere could provide a social argumentation environment for wider experimentation and collaboration.

Acknowledgements: Cohere is implemented by Michelle Bachler, to whom I am indebted. The Open Sensemaking Communities Project [http://kmi.open.ac.uk/projects/osc] is funded by the William and Flora Hewlett Foundation, as part of the Open University’s OpenLearn Initiative: www.open.ac.uk/openlearn

6. References
