Multiuser Museum Interactives for Shared Cultural Experiences: an Agent-based Approach

Matthew Yee-King
Goldsmiths College
University of London
London SE14 6NW, UK
mas01mjy@gold.ac.uk

Roberto Confalonieri
Goldsmiths College
University of London
London SE14 6NW, UK
confalonieri@lsi.upc.edu

Dave de Jonge
Artificial Intelligence Research
Institute (CSIC-IIIA)
08193 Bellaterra, Spain
davedejonge@iiia.csic.es

Katina Hazelden
Goldsmiths College
University of London
London SE14 6NW, UK
exs01kh2@gold.ac.uk

Carles Sierra
Artificial Intelligence Research
Institute (CSIC-IIIA)
08193 Bellaterra, Spain
sierra@iiia.csic.es

Mark d'Inverno
Goldsmiths College
University of London
London SE14 6NW, UK
dinverno@gold.ac.uk

ABSTRACT

Multiuser museum interactives are computer systems installed in museums or galleries which allow several visitors to interact together with digital representations of artefacts and information from the museum's collection. WeCurate is such a system, providing a multiuser curation workflow where the aim is for the users to synchronously view and discuss a selection of images, finally choosing a subset of these images that the group would like to add to their group collection. The system presents two main problems: workflow control and group decision making. An Electronic Institution (EI) is used to model the workflow into scenes, where users engage in specific activities in specific scenes. A multiagent system is used to support group decision making, representing the actions of the users within the EI, where the agents advocate and support the desires of their users e.g. aggregating opinions, proposing interactions and resolutions between disagreeing group members and choosing images for discussion.

Categories and Subject Descriptors

I.2 [**Artificial Intelligence**]: Distibuted Artificial Intelligence—*Intelligent agents, Multiagent systems*

General Terms

Algorithms, Experimentation

Keywords

Multiagent system, Collective decision making

1. INTRODUCTION

In recent times, high tech museum interactives have become ubiquitous in major institutions. Typical examples include augmented reality systems, multitouch table tops and virtual reality tours [9, 12, 21]. Whilst multiuser systems have begun to appear, e.g. a 10 user quiz game in the Tate Modern, the majority of these museum interactives do not perhaps facilitate the sociocultural experience of visiting a museum with friends, often being designed for a single user. At this point, we should note that mediating and reporting the actions of several 'agents' to provide a meaningful and satisfying sociocultural experience for all is challenging, requiring multiple criteria decision making. Another trend in museum curation is the idea of community curation, where a community discourse is built up around the artefacts, to provide different perspectives and insights [20].[1] This trend is not typically represented in the design of museum interactives, where information-*browsing*, not information-*generation* is the focus. However, museums are engaging with the idea of crowdsourcing with projects such as 'Your Paintings Tagger' and 'The Art Of Video Games' [10, 4]. Again, controlling the workflow within a group to engender discussion and engagement with the artefacts is challenging, especially when the users are casual ones as in a museum context.

In this paper we propose a first of its kind multiuser museum interactive which uses a multiagent system to support community interactions and decision making and an electronic institution to model the workflow. Our aim is not only to make use of agent technology and electronic institutions as a means to implement a multiuser museum interactive, but also to relate agent theory to practice in order to create a usable system that will help us to answer the following research questions. Do agent technology and electronic institutions enable users to share online experiences in a way that was not possible before? Are decision making capabilities and aggregation operators defined in the literature suitable for enhancing the experience of users?

To this end, we specify a community curation session in terms of the scenes of an electronic institution for controlling community interactions. We support system's and users' decisions by means of personal assistant agents equipped with different decision making capabilities. We make use of a

Appears in: *Proceedings of the 12th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2013), Ito, Jonker, Gini, and Shehory (eds.), May, 6–10, 2013, Saint Paul, Minnesota, USA.*

[1] Edmond de Goncourt's opinion that 'A painting in a museum hears more ridiculous opinions than anything else in the world' not withstanding.

multimodal user interface which directly represents users as agents in the scenes of an underlying electronic institution and which is designed to engage casual users in a social discourse around museum artefacts. We describe a method for evaluating such systems which combines data driven analysis with user feedback.

The system described here is an improvement of a previous version trialled in May 2012. This second version is due to be installed for two weeks at a major London museum in November 2012 with the plan of creating a permanent exhibit and then moving the technology onto other museums.

The paper is organised as follows. Section 2 presents several works that relate to ours from different perspectives. Section 3 describes the WeCurate workflow and the underlying system. Section 4 discusses some preliminary results we obtained in the trials of a previous version of the system and outlines the evaluation scheme we aim to use to validate the current version. Finally, Section 5 draws some conclusions and envisions some of the ideas we have in mind to improve the current system.

2. RELATED WORK

Many systems exist that enable realtime personalised experience and multiuser collaboration in virtual workspaces, both in industry and in academia. In industry, web conferencing software such as Adobe Connect allows complex media and text driven interactions; shared document editors such as Google Drive enable co-editing of office-type documents. However, the user interfaces are perhaps too complex for a casual user in a museum and it is not possible to enforce specific workflows with specific goals with these systems as required by our group curation scenario. Further, agreement technologies such as group decision making are not explicitly supported, e.g. consider the scenario where users are co-editing a presentation using Google Drive and they need to select an appropriate image.

In academia, enhancing the users' experience in museums has already been addressed in different ways. For instance, in the PEACH project, researchers focused on the creation of online personalised presentations to be delivered to the visitors for improving their satisfaction and personalised visit summary reports of suggestions for future visits [13]. Their focus was mainly the modeling of preferences of single users but the importance of social interactions in visiting a museum was investigated in the PIL project, an extension of the research results of the PEACH project, and in the ARCHIE project [14, 15]. ARCHIE aimed to provide a more socially-aware experience to users visiting a museum by allowing visitors to interact with other visitors by means of their mobile guides. User profiles were used to tailor the information to the needs and interests of each individual user and, as such, no group decision making was necessary. A cultural heritage application was proposed in [6] where agents are able to discover users' movements via a satellite, to learn and to adapt user profiles to assist users during their visits in Villa Adriana, an archaeological site in Tivoli, Italy.

On the other hand, in a system aiming to create a multiuser museum interactive, it is not only important to model users (by means of agents) and their preferences but also to assist them in taking decisions. For example, the system could decide which artefact is worthy to be considered by merging users' preferences [23]; or it could decide whether the artefact is collectively accepted by a group of users by considering users' valuations about certain criteria of the artefact itself like in multiple criteria decision making [17]; or assist users in reaching agreements by arguments' exchange like in argument-based negotiation [1]. These cases, that correspond to different decision making problems can be solved by defining different decision principles which take the preferences of the single users into account and compute the decision of the group as a whole. One way to define a decision principle is by means of an aggregation operator. Several operators for aggregating data have been proposed in the literature [3, 22, 19].

In [3], a bipolar fusion operator for merging users' preferences has been studied in the possibilistic logic setting. According to this approach, the problem of deciding what is collectively accepted can be handled by means of an aggregation function on the whole set of positive and negative preferences (represented in terms of possibility distributions) of a group of agents. The Ordered Weighted Average (OWA) operator [22] allows to weight values in relation to their ordering. In this way, a system can give more importance to a subset of the input values than to another subset, e.g. extreme values versus central ones. A generalisation of both the weighted mean and the OWA operator is the Weighted Ordered Weighted Average (WOWA) operator [19]. WOWA allows the system to weight the importance of a value (as in the weighted mean), and the values in relation to their relative position (as in the OWA operator). The choice of the operator to be used usually depends on the type of experience the system aims to provide to its users.

A multiuser interactive is a typical example of a system in which users (and thus their respective agents) can enter and leave the system and they behave according to *norms* that are appropriate for that specific society. A convenient way to model social interactions and their norms is by means of an Electronic Institution (EI) [2].

An EI makes it possible to develop programs according to a new paradigm, in which the tasks are executed by independent agents, that are not necessarily designed specifically for the given program and that cannot be blindly trusted. An EI is responsible for making sure that the agents behave according to the *norms* that are necessary for the application. Therefore, the EI paradigm allows a flexible and dynamic infrastructure, in which agents can interact in an autonomous way within the norms of the cultural institution.

EIs have usually been considered as centralised systems [16, 8, 7]. Nevertheless, since users can be physically in different places, it is desirable to run an EI in a distributed manner to enable a flexible and dynamic structure, which is characteristic to human social communities.

3. IMPLEMENTATION

WeCurate is a museum interactive which provides a multiuser curation workflow where the aim is for the users to synchronously view and discuss a selection of images, finally choosing a subset of these images that the group would like to add to their group collection. In the process of curating this collection, the users are encouraged to develop a discourse about the images in the form of weighted tags and comments as well as a process of bilateral argumentation. Further insight into user preferences and goals is gained from data about specific user actions such as image zooming and general activity levels. The activity has been defined as an electronic institution, with agents operating

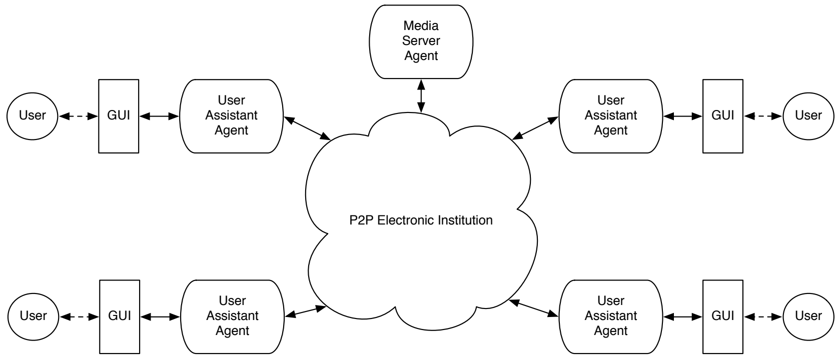

Figure 1: The WeCurate system architecture, here showing 4 simultaneous users.

within it to represent the activities of the users and to provide other services. The users interact with the system using an animated user interface. An overview of the system architecture, showing the electronic institution, agent and user interface components is provided in figure 1.

In the following sections, we deal first with the way we have modelled the WeCurate workflow using an electronic institution. Then we describe the agents that enact the workflow, with particular emphasis on user representation and group decision making. The user interface is presented, with images of the different scenes in the workflow. The system architecture is described, including the connections between EI, agents and UI. Finally, the adopted technologies used to implement the system are briefly explained.

3.1 WeCurate workflow

The WeCurate workflow is defined as an EI consisting of 5 scenes, with associated rules controlling messaging and transitions between scenes. An overview of the workflow is provided in figure 2. The scenes are as follows:

- **Login and lobby scene**: this allows users to login and wait for other users to join. The EI can be configured to require a certain number of users to login before transition to the selection scene can take place.

- **Selection scene**: its purpose is to allow a quick decision as to whether an image is interesting enough for a full discussion. Users can *zoom* into the image and see the zooming actions of other users. They can also set their *overall preference* for the image using a like/dislike slider. The UI is shown in figure 3a.

- **Forum scene**: if an image is deemed interesting enough, the users are taken to the forum scene where they can engage in a full *discussion* of the image. Users can add and delete tags, they can resize tags to define their opinions of that aspect of the image, they can make comments, they can zoom into the image and they can see the actions of the other users. They can also view images that were previously added to the collection and choose to argue with another user directly. The aim is to collect community information about the image. The UI is shown in figure 3b.

- **Argue scene**: here, two users can engage in a process of *bilateral argumentation*, wherein they can propose aspects of the image which they like or dislike, in the form of tags. The aim is to convince the other user to align their opinions with yours, in terms of tag sizes. For example, one user might like the 'black and white' aspect of an image, whereas the other user dislikes it; one user can then pass this tag to the other user to request that they resize it. The UI is shown in figure 3c.

- **Vote scene**: here, the decision is made to add an image to the group collection or not by *vote*. The UI is shown in figure 3d.

3.2 Institutional and other agents

This electronic institution itself is executed by several institutional agents, including a *SceneManager* which runs the scene instances, an *EIManager* which admits agents to the EI and instantiates scenes and several *Governors* which control message passing between agents. This first non-institutional agent is the *MediaAgent*, which provides access to the media archive. The users are represented by *UserAssistant* agents. Each time a new user joins a curation session, a new UserAssistant is enrolled into the EI. The actions of the users are passed down to their representative agents which then send messages to each other via the EI. An agent's Governor will check if a given message can be sent, based on the state of the current scene and the role of the sender. The state of a scene might change as a result of a message being sent. For example, a user clicking on the 'go to forum' button in the selection scene will generate a message which will trigger a count down timer at the SceneManager level, giving the other users a limited time in which to complete their examination of the image before a scene transition occurs. Other messages might not affect the state of the scene, but instead will provide information to other agents about the actions of a user, e.g. zooming into an image or changing a tag preference. This information might be used later on to make decisions, for example.

In the following section, the decision making criteria used in the WeCurate workflow are described.

3.3 Agent supported decision making

The UserAssistant agents are involved in several automatic decisions and calculations. These are described in the following subsections.

3.3.1 Selection scene

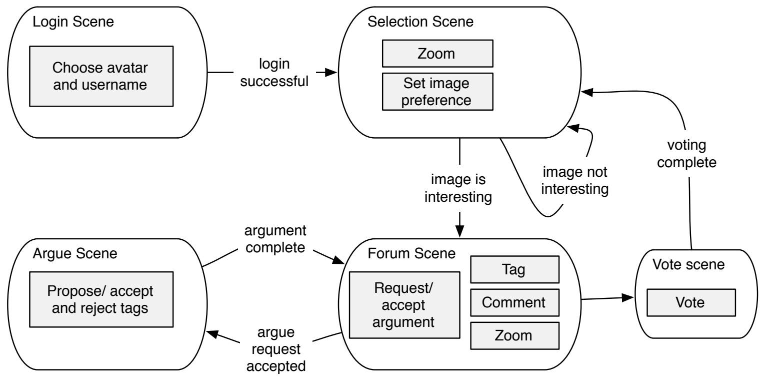

Figure 2: The WeCurate workflow: white boxes represent scenes, grey boxes represent users' actions, and arrows denote scenes' transitions.

In the selection scene, if the image is deemed sufficiently interesting, the agents (and therefore their users) are taken to the forum scene. *Image interestingness* of an image takes the image preferences of all the users into account and it is computed in different ways.

Image interestingness based on average.

The *image preference* of a user is a value belonging to a finite bipolar scale $\mathcal{S} = [-1, -0.9, \ldots, +0.9, +1]$ where -1 stands for "discard" and $+1$ stands for "keep". Given a group of users $\{u_1, u_2, \ldots, u_n\}$, let us denote the image preference of a user u_i w.r.t an image im by $r_i(im) = v_i$ with $v_i \in \mathcal{S}$.

Each UserAssistant agent also stores the zoom activity of its user. Clearly, the zoom activity is a measure of the user's interest in a given image and, as such, it should be taken into account in the calculation of the image interestingness. Let us denote the number of *image's zooms* of user u_i w.r.t. an image im as $z_i(im)$. Then, we can define the total number of zooms for an image im as $z(im) = \sum_{1 \le i \le n} z_i(im)$. Based on $z(im)$ and the z_i's associated with each user, we can define a weight for the image preference r_i of user u_i as $w_i = z_i/z(im)$.

The *image interestingness* of n users w.r.t an image im, denoted by $\overline{r(im)}$, can be defined as:

$$\overline{r(im)} = \frac{\sum_{1 \le i \le n} r_i w_i}{\sum_{1 \le i \le n} w_i} \tag{1}$$

Then, a decision criterion for the interestingness of an image im is defined as follows:

$$\texttt{int}(im) = \begin{cases} 1, & \text{if } 0 \le \overline{r(im)} \le 1 \\ 0, & \text{if } -1 \le \overline{r(im)} < 0 \end{cases} \tag{2}$$

Therefore, the system proceeds with a forum scene when $\texttt{int}(im) = 1$, while the system goes back to a select scene when $\texttt{int}(im) = 0$.

Image interestingness based on WOWA operator.

An alternative criterion for deciding whether an image is interesting or not can be defined by using a richer average operator such the Weighted Ordered Weighted Average (WOWA) operator [19].

The WOWA operator is an aggregation operator which allows to combine some values according to two types of weights: i) a weight referring to the importance of a value itself (as in the weighted mean), and ii) an ordering weight referring to the values' order. Indeed, WOWA generalizes both the weighted average and the ordered weighted average [22]. Formally, WOWA is defined as [19]:

$$f_{wowa}(r_1, \ldots, r_n) = \sum_{1 \le i \le n} \omega_i r_{\sigma(i)} \tag{3}$$

where $\sigma(i)$ is a permutation of $\{1, \ldots, n\}$ such that $r_{\sigma(i-1)} \ge r_{\sigma(i)}$ $\forall i = 2, \ldots, n$, ω_i is calculated by means of an increasing monotone function $w^*(\sum_{j \le i} p_{\sigma(j)}) - w^*(\sum_{j < i} p_{\sigma(j)})$, and $p_i, w_i \in [0, 1]$ are the weights and the ordering weights associated with the values respectively (with the constraints $\sum_{1 \le i \le n} p_i = 1$ and $\sum_{1 \le i \le n} w_i = 1$).

We use the WOWA operator for deciding whether an image is interesting in the following way. Let us take the weight p_i for the image preference r_i of user u_i as the percentage of zooms made by the user (like above). As far as the ordering weights are concerned, we can decide to give more importance to image preference's values closer to extreme value such as -1 and $+1$, since it is likely that such values can trigger more discussions among the users rather than image preference' values which are close to 0. Let us denote the sum of the values in $\mathcal{S}$ as s. Then, for each image preference $r_i(im) = v_i$ we can define an ordering weight as $w_i = r_i(im)/s$. Please notice how the p_i's and w_i's defined satisfy the constraints $\sum_{1 \le i \le n} p_i = 1$ and $\sum_{1 \le i \le n} w_i = 1$.

Then, a decision criterion for the interestingness of an image im based on $f_{wowa}(r_1, \ldots, r_n)$ can be defined as:

$$\texttt{int}_{\texttt{wowa}}(im) = \begin{cases} 1, & \text{if } 0 \le f_{wowa}(r_1, \ldots, r_n) \le 1 \\ 0, & \text{if } -1 \le f_{wowa}(r_1, \ldots, r_n) < 0 \end{cases} \tag{4}$$

3.3.2 *Forum Scene*

The main goal of the users in a forum scene is to discuss an image, which has been considered interesting enough in a select scene, by pointing out what they like or dislike of the image through tags. During the tagging, the *overall image preference per user* is automatically updated. Whilst tagging is the main activity of this scene, a user can also choose to argue with another user in order to persuade him to adopt his own view (i.e. to "keep" or to "discard" the image). In such a case, a list of *recommended argue candidates* is retrieved. Finally, when a user is tired of tagging, he can propose the other users to move to a vote scene. In this case, an automatic *multi-criteria decision* is taken in order to decide whether the current image can be added or not to the image collection without a vote being necessary.

Overall image preference per user.

In order to update the overall image preference of a user user in an automatic way, all the tags which the user has been specifying w.r.t. an image must be taken into account. An *image's tag* (or image's argument) t for an image im is represented as $(\langle id, v, w\rangle, im)$ where id identifies a certain feature of the image, such as "blue", "sea" etc., $v \in \mathcal{S}$ is the tag's value and $w \in \{0, 0.1, \ldots, 0.9, 1\}$ is the tag's importance.

Let $\mathcal{T}^i_{im} = \{t_1, t_2, \ldots, t_m\}$ be the set of image's tags specified by a user u_i w.r.t. an image im. Then, the *aggregate* or *overall image preference* of a user u_i given $\mathcal{T}^i_{im}$ denoted by $r^*_i(im)$ is defined as:

$$r^*_i(im) = \frac{\sum_{1\le j\le m} v_j w_j}{\sum_{1\le j\le m} w_j} \tag{5}$$

When a user rates the image im by specifying of a new tag or by updating a tag already specified, his overall image preference is automatically updated by computing $r^*_i(im)$.

Argue Candidate Recommender.

In order to recommend an ordered list of argue candidates to a user willing to argue, the distance between the overall image preferences per user (Eq.5) can be taken into account.

Let u_i be a user willing to argue and $r^*_i(im)$ be his overall image preference. Then, for each u_j (such that $j \neq i$) we can define the *image preference distance* of user u_j w.r.t. user u_i, denoted by $\delta_{ji}(im)$, as:

$$\delta_{ji}(im) = \{\mathtt{abs}(r^*_j(im) - r^*_i(im)) | (r^*_j(im) < 0 \wedge r^*_i(im) \ge 0) \vee (r^*_j(im) \ge 0 \wedge r^*_i(im) < 0)\} \tag{6}$$

Then, an argue candidate for user u_i for an image im is $cand_i(im) = \{u_j \mid \max\{\delta_{ji}(im)\}\}$. The ordered list of argue candidates can be defined by ordering the $\delta_{ji}(im)$'s.

Multi-criteria decision.

When users in a forum scene decide to move to a vote scene, first, an automatic decision process to decide whether the image can be added or not to the image collection is run. The problem of deciding whether to add the image is a *multi-criteria decision problem* in which different users can have different criteria (their image's tags). The main idea behind this decision step is to first obtain the overall image preference of each user, and, then, to take the image's preferences into account in order to define a decision criterion.

The overall image preference of each user can be computed according to Eq.5. Then, a decision criterion for deciding whether to add an image im or not to the image collection can be defined as:

$$\mathtt{add}(im) = \begin{cases} 1, \text{ if } \forall u_i, 0 \le r^*_i(im) \le 1 \\ 0, \text{ if } \forall u_i, -1 \le r^*_i(im) < 0 \\ undefined, \text{ otherwise} \end{cases} \tag{7}$$

Therefore, an image im is automatically added to the image collection if it has been unanimously accepted by the users. On the contrary, the image is discarded if it has been unanimously rejected. Finally, if $\mathtt{add}(im) = undefined$, then the system is unable to decide and the final decision is taken by the users in a vote scene.

3.3.3 Argue Scene

The main goal of two users running in an argue scene is to try to reach an agreement about to "keep" or to "discard" an image im by exchanging image's tags (image's arguments). The argue scene defines a bilateral argumentation protocol. The formal protocol was presented in [1] and it works as follows:

- the two users tag the image im by means of image's tags (like in the forum scene), but, they can also propose their image's tags to the other user:
  - while tagging, their overall image preferences are automatically updated;
- a user proposes an image's tag to the other user who can either accept or reject it:
  - if the user accepts the image's tag proposed, then their overall image preferences are automatically updated:
    - if an *argue agreement* is reached, then the argue scene stops,
    - otherwise, the argue scene keeps on;
  - if the user rejects the image's tag proposed, then the argue scene keeps on;

Both users can also decide to leave the argue scene spontaneously.

Overall image preference per user.

The overall image preference of a user in an argue scene is automatically updated by computing $r^*(im)$ (see subsection 3.3.2).

Argue Agreement.

Informally, an *argue agreement* is reached when the image preferences of the two users agree towards "keep" or "discard". Let $r^*_i(im)$ and $r^*_j(im)$ be the image's preferences of user u_i and u_j respectively. Then, a decision criterion for deciding whether an argue agreement is reached can be defined as:

$$\mathtt{argue}(im) = \begin{cases} 1, \text{ if } (0 \le r^*_i(im) \le 1 \wedge 0 \le r^*_j(im) \le 1) \vee \\ (-1 \le r^*_i(im) < 0 \wedge -1 \le r^*_j(im) < 0) \\ 0, \text{ otherwise} \end{cases} \tag{8}$$

Therefore, an argue scene stops when $\mathtt{argue}(im) = 1$. Instead, while $\mathtt{argue}(im) = 0$, the argue scene keeps on until either $\mathtt{argue}(im) = 1$ or the two users decide to stop arguing.

3.3.4 Vote Scene

The main goal of the users running in a vote scene is to decide by vote to add or not an image to the image collection. This decision step occurs when the automatic decision process at the end of the forum scene is unable to make a decision.

In a vote scene, each user's vote can be "yes", "no", or "abstain" (in case that no vote is provided). Let $v_i \in \{+1, 0, -1\}$ be the vote of user u_i where $+1$ = "yes", -1 = "no", and 0 = "abstain" and let $\mathcal{V} = \{v_1, v_2, \ldots, v_n\}$ be the set of votes of the users in a vote scene. Then, a decision criterion for adding an image or not based on *vote counting* can be defined as:

$$\texttt{vote}(im) = \begin{cases} 1, & \text{if } \sum_{1 \leq i \leq n} v_i \geq 0 \\ 0, & \text{otherwise} \end{cases} \quad (9)$$

Therefore, an image im is added to the image collection if the number of "yes" is greater or equals than the number of "no". In the above criterion, a neutral situation is considered as a positive vote.[2]

3.4 User interface

The user interface provides a distinct screen for each scene, as illustrated in figures 3a, 3b, 3c and 3d. It communicates with the UserAssistant agent by sending a variety of user triggered events which are different in each scene. The available user actions in each scene are shown in figure 1. The state of the interface is completely controlled by the UserAssistant agents, which send scene snapshots to the interface whenever necessary, e.g. when a new tag is created. Some low level details of the method of data exchange between interface and agents are provided in the adopted technologies section below (3.5).

The interface is the second iteration of a shared image browsing interface, designed to include desirable features highlighted by a user trial of the first iteration [11]. Desirable features include standard usability such as reliability, speed and efficiency etc., awareness of the social presence of other users and awareness of the underlying workflow. Given the social nature of the system, social presence, where users are aware of each others' presence and actions as well as a shared purpose and shared synchronicity is of especial interest. The detailed concepts behind the design are beyond the scope of this paper and will be reported elsewhere.

3.5 Adopted technologies

The physical museum installation includes 4 iPads allowing 4 simultaneous users to use the system. There is a large panel displaying a summary of what is going on or information about the exhibit, during inactive periods. A local network connects the iPads to a local server, which executes the agents and EI, as well as driving the information screen. The institutional and other agents are implemented as independent Java programs and they discover and communicate with each other and the components of the EI using a FreePastry P2P network[18]. This means the agents and EI system can be run in a distributed fashion if necessary. The user interface is implemented using Javascript drawing to an HTML5 canvas element, which is a cross platform and plugin free solution. The downside is that the interface cannot communicate directly with the agents since it is not a part of the FreePastry network. Instead, the interface sends events formatted as JSON to a queue stored on an HTTP server. The agents poll the server and pick up the event queue, then in turn generate scene snapshots in JSON format which are sent to the HTTP server. The interface polls the HTTP server for the latest snapshot, which is then used to define the state of the interface.

One advantage of this queued event and snapshot model with regard to evaluation is that all interface events and interface state snapshots are stored on the server for later inspection. This allows a complete, interactive reconstruction of activity of the users and the agents for qualitative analysis as well as providing a lot of data for quantitative analysis. The evaluation scheme section (4) describes how this data will be used.

4. EVALUATION

The objective of the evaluation was twofold. First, to determine how WeCurate would be used by social groups located onsite at the museum; second, to understand whether our agent decision models could predict which images users would like from users' preferences.

4.1 Trial and Study

The WeCurate system was set up as an interactive exhibit in a major London museum, supported by the research team. Up to four visitors used WeCurate running on 4 iPads, fixed around a single table so the users were co-located. The museum provided 150 images from their collection; the task for the group was to decide which of these images they would like to have as a postcard, via the curation process.

Multiple sources of qualitative and quantitative data were collected. Participants were filmed during the activities and their interactions with the system was recorded in timestamped logs. Data was gathered and cross referenced from adhoc observation of the trials themselves, inspection of the video footage, transcription of the interviews and the system log files.

The ages of participants ranged from 4 years (with assistance) to 45 years. The average time each group used the WeCurate system was 5 mins 38 secs, the longest session logged was 21 mins 16 seconds. 224 observations about images' evaluations were collected. According to these observations, 130 images were chosen from the original set and the 73,1% was finally added to the group collection. Each image was seen from 1 to 4 times during the different sessions. Among the 224 observations, 176 corresponded to image evaluations in which an image was added to the group collection and 48 in which an image was rejected by voting.

Of particular interest was the communication and discussion about the artefacts presented by the system, and whether the shared view supported an awareness of social action. The observational data was coded using a grounded theory approach where a code schema emerges from behaviours and opinions as opposed to being pre-ordained [5]. The observations also provided a dataset for assessing the success of the different types of agent decision models. To this end, we compared the image interestingness and the multiple criteria measures calculated by the agents in the selection and forum scenes w.r.t. the users' decisions in the vote scene.

4.2 Results

Dynamic between adults and between parents and children: Of the adult only sessions, 70% featured some degree of laughter and playful comments, these included reactions to

[2]This assumption is made to avoid an *undecided* outcome in this decision step.

(a) The selection scene for rapid selection of interesting images

(b) The forum scene for in-depth discussion

(c) The argue scene for bilateral argumentation

(d) The vote scene for deciding to add images to the group collection

Figure 3: The WeCurate user interface. Bubbles represent tags and are resizable and movable; icons visible on sliders and images represent users.

another participant deleting a newly created tag, or commenting on the content of a tag. Consequently for the adults, the creation of a tag, or modifying a group member's tag was often perceived as a playful action. 60% of the adult's sessions also featured an attempt by at least one of the participants to synchronise their actions with the group (i.e. not clicking "next" until others were ready to move to the next image / vote). Aside from the positive communication among the adults, there were instances in 60% of these sessions where a participant expressed an opinion or asked for an opinion and no one responded. The lack of acknowledgement of group members comments could indicate that the participants were too engaged with the task and therefore did not register the comment, or they simply chose to ignore the group member. The social dynamic between parent and child was dominated by adult initiated action whereby 89% of the interactions related to the adult driving the child's comprehension. Of the adult initiated behaviour, 40% was directing the child's action and attention, and 45% was requesting an opinion about the image from the child.

Discussion of task, image and museum artefacts: The questionnaire showed that 56% reported feeling as if they had a full discussion, while 23% reported that they did not (21% did not comment). Whilst it is encouraging that a majority believed they had a rich debate about the images in the system, as this a key aspect of the design and use, a more significant margin would be preferable. Of more concern is that in 30% of the sessions observed (with both adults and children) there was no discussion between the participants using separate devices, and in only one of these sessions did the children talk to each other (in all other sessions they conversed solely with their parent). The absence of discussion could be partially accounted for by the parents preoccupation with supporting their child.

Social awareness via the system: When reporting on their ability to express an opinion of the image in the questionnaire, 73% of participants felt they were able to express a preference in the selection scene, and 81% reported that they could express opinions via the forum scene using the tags. This suggests that the participants felt they were able to communicate their preferences via the WeCurate Interface. The social group did appear to have some influence over individual's decision making, whereby 42% reported changing their decision as a consequence of seeing other's actions.

Agents' decision capabilities: From the observation data, the average and WOWA operators were respectively classifying as interesting the 53% and the 26% of the images which were finally positively voted, and as not interesting the 56% and the 67% of the images that were finally discarded. These results show that the average and WOWA operators were too weak and too strong at the moment of selecting the images. This suggests that a combination of these operators can enhance the system's prediction. For instance, by using the average operator to select images, but to finally vote only those images which are not discarded by the WOWA. Concerning the multiple criteria operator, we observed that 73% of the images chosen in the forum was positively voted and 21% of them was correctly filtered, that is, in only few forum evaluations, images were classified as not worthy to be added to the group collection before the vote. This can suggest that either the users were too conservative about their opinions (created in the selection scene), or that the process of opinions' revision by means of weighted tags was too complex for the users.

5. CONCLUSION AND FUTURE WORK

A first of its kind multiuser museum interactive which uses a multiagent system to support community interactions and decision making and an Electronic Institution (EI) to model the workflow has been described. Its multimodal user interface which directly represents the scenes in an underlying EI and which is designed to engage casual users in a social discourse around museum artefacts has also been described. An evaluation of the users' social dynamics and of the collective decision models of the agents has been presented.

This line of research looks promising. Our results have shown that the representations of the opinions of the group did influence individual members opinion, which denotes a sense of social presence via the system. Although from the evaluation it has emerged that the decision models employed by the agents, if taken individually, tend not to predict users' final decisions, and that an alternative design of the discussion scene can be valuable to encourage users to revise their opinions, the use of agent technology and EIs enhanced users' social dynamics and users' social presence.

Based on these results, the design of the system at all levels will undergo a further iteration, with an aim to improving users' social engagement, the scenes' design and, the efficacy of the agent architecture in supporting the curation task. For instance, to let users be more engage in the discussion by designing the forum scene built around competitiveness rather than cooperativeness and to enhance the prediction capabilities of the system by combining the agents' decision models. Beyond that, the technology has been designed to easily transfer to a web or mobile app, and the distributed agent execution model is designed to scale; and we see great potential in the concept of agent supported, workflow driven, synchronous image discussion and curation taken to the mass audience on the open web.

6. ADDITIONAL AUTHORS

Additional authors: Leila Amgoud (IRIT, Universitè Paul Sabatier, 118 Route de Narbonne, 31062 Toulouse, France, email: amgoud@irit.fr) and Nardine Osman (CSIC-IIIA, 08193 Bellaterra, Spain, email: nardine@iiia.csic.es).

7. REFERENCES

[1] L. Amgoud, R. Confalonieri, D. D. Jonge, M. d'Inverno, K. Hazelden, N. Osman, H. Prade, C. Sierra, and M. Yee-King. Sharing Online Cultural Experiences: An Argument-Based Approach. In V. Torra, *et. al.*, editors, *Modeling Decisions for Artificial Intelligence - 9th Int. Conf., MDAI 2012*, vol. 7647 of *LNCS*, pages 282–293. Springer, 2012.

[2] J. L. Arcos, M. Esteva, P. Noriega, J. A. Rodríguez-Aguilar, and C. Sierra. Engineering open environments with electronic institutions. *Eng. Applications of Art. Int.*, 18(2):191–204, 2005.

[3] S. Benferhat, D. Dubois, S. Kaci, and H. Prade. Bipolar possibility theory in preference modeling: Representation, fusion and optimal solutions. *Information Fusion*, 7(1):135–150, 2006.

[4] C. Barron, The Washington Post. The Art of Video Games, Apr. 2012.

[5] K. Charmaz. *Constructing grounded theory: A practical guide through qualitative analysis.* Sage Publications Limited, 2006.

[6] S. Costantini, L. Mostarda, A. Tocchio, and P. Tsintza. DALICA: Agent-Based Ambient Intelligence for Cultural-Heritage Scenarios. *IEEE Intelligent Systems*, 23(2):34–41, Mar. 2008.

[7] M. d'Inverno, M. Luck, P. Noriega, J. A. Rodríguez-Aguilar, and C. Sierra. Communicating open systems. *Artificial Intelligence*, 186:38–94, 2012.

[8] M. Esteva. *Electronic Institutions: From Specification to Development.* PhD thesis, Technical University of Catalonia, 2003.

[9] A. Gaitatzes and M. Roussou. Reviving the past : Cultural Heritage meets. In *Proc. of the 2001 conference on Virtual reality, archeology, and cultural heritage*, pages 103–110, 2002.

[10] A. Greg. Your paintings: public access and public tagging. *Journal of the Scottish Society for Art History*, 16, 2011.

[11] K. Hazelden, M. Yee-King, L. Amgoud, M. d'Inverno, C. Sierra, N. Osman, R. Confalonieri, and D. D. Jonge. WeCurate: Designing For Synchronised Browsing And Social Negotiation. In *Proc. of the 1st Int. Conf. on Agreement Technologies, AT 2012*, pages 168–179, 2012.

[12] E. Hornecker. "I don't understand it either, but it is cool" - visitor interactions with a multi-touch table in a museum. In *2008 3rd IEEE Int. Workshop on Horizontal Interactive Human Computer Systems*, pages 113–120, 2008.

[13] T. Kuflik, C. Callaway, D. Goren-Bar, C. Rocchi, O. Stock, and M. Zancanaro. Non-intrusive user modeling for a multimedia museum visitors guide system. In L. Ardissono, *et. al.*, editors, *Proc. of the 10th Int. Conf. on User Modeling*, vol. 3538 of *LNCS*, pages 236–240. Springer, 2005.

[14] T. Kuflik, O. Stock, M. Zancanaro, A. Gorfinkel, S. Jbara, S. Kats, J. Sheidin, and N. Kashtan. A visitor's guide in an active museum: Presentations, communications, and reflection. *J. Comput. Cult. Herit.*, 3(3):11:1–11:25, Feb. 2011.

[15] K. Luyten, H. van Loon, D. Teunkens, K. Gabriels, K. Coninx, and E. Manshoven. ARCHIE: Disclosing a Museum by a Socially-aware Mobile Guide. In Proc. of the 7th Int. Symp. on Virtual Reality, Archaeology and Cultural Heritage (VAST 2006), 2006.

[16] P. Noriega. *Agent Mediated Auctions: The Fishmarket Metaphor.* PhD thesis, Autonomous University of Barcelona, 1997.

[17] R. A. Ribeiro. Fuzzy multiple attribute decision making: a review and new preference elicitation techniques. *Fuzzy Sets Syst.*, 78(2):155–181, Mar. 1996.

[18] A. Rowstron and P. Druschel. Pastry: Scalable, decentralized object location, and routing for large-scale peer-to-peer systems. In *Middleware 2001.*

[19] V. Torra. The weighted owa operator. *International Journal of Intelligent Systems*, 12(2):153–166, 1997.

[20] C. Turner. Making Native Space: Cultural Politics, Historical Narrative, and Community Curation at the National Museum of the American Indian. *Practicing Anthropology*, 33(2):40–44, Apr. 2011.

[21] R. Wojciechowski, K. Walczak, M. White, and W. Cellary. Building virtual and augmented reality museum exhibitions. In *Proc. of the 9th Int. Conf. on 3D Web technology*, pages 135–144. ACM, 2004.

[22] R. R. Yager. On ordered weighted averaging aggregation operators in multicriteria decisionmaking. *IEEE Trans. Syst. Man Cybern.*, 18(1):183–190, 1988.

[23] M. Yee-King, R. Confalonieri, D. D. Jonge, N. Osman, K. Hazelden, L. Amgoud, H. Prade, C. Sierra, and M. d'Inverno. Towards Community Browsing for Shared Experiences: The WeBrowse System. In *Proc. of the 1st Int. Conf. on Agreement Technologies, AT 2012*, pages 201–202, 2012.