

# Network Arts: Exposing Cultural Reality

David A. Shamma\*, Sara Owsley\*, Kristian J. Hammond\*,  
Shannon Bradshaw†, Jay Budzik\*

\*Intelligent Information Laboratory  
Northwestern University  
1890 Maple Avenue, 3rd Floor  
Evanston, Illinois 60201  
+1 847 467 1771

{ayman, sowsley, hammond, budzik}@cs.northwestern.edu

†Department of Management Sciences  
The University of Iowa  
Iowa City, Iowa 52242  
+1 319 335 3944  
shannon-bradshaw@uiowa.edu

## ABSTRACT

In this article, we explore a new role for the computer in art as a reflector of popular culture. Moving away from the static audio-visual installations of other artistic endeavors and from the traditional role of the machine as a computational tool, we fuse art and the Internet to expose cultural connections people draw implicitly but rarely consider directly. We describe several art installations that use the World Wide Web as a reflection of cultural reality to highlight and explore the relations between ideas that compose the fabric of our every day lives.

## Categories and Subject Descriptors

J.5 [Arts and Humanities]: Fine arts; H.3.m [Information Storage and Retrieval]: Miscellaneous; H.5.3 [Information Interfaces and Presentation]: Group and Organization Interfaces—*Web-based interaction*

## General Terms

Human Factors

## Keywords

Network Arts, Media Arts, Culture, World Wide Web, Information Retrieval, Agents

## 1. INTRODUCTION

The Web has evolved to play many roles in our lives. One of the more interesting, yet unexploited, is its role as a storehouse of cultural connections; portals, blogs, and other types of sites are a reflection of popular culture. We have created a set of systems that expose and highlight the connections people use on a daily basis, but rarely consider. These systems, by making their processes visible elevate the mundane, the available, and the purely popular. As

Copyright is held by the author/owner(s).  
WWW2004, May 17–22, 2004, New York, NY USA.  
ACM xxx.xxx.



Figure 1: The Imagination Environment running a performance on the wall while watching a DVD of *The Godfather*.

“artistic agents” they gather, sift, and present our reality back to us as they move through networks of information.

### 1.1 Artistic Focus

Our work in this arena over the past two years has resulted in an exciting set of installations. Each has its own dynamic; each its own deployment. Each has its own way of using the Web to give the piece its own force. Though very different, each installation was created to expose the power of the Web as a reflector of our broad and diverse global culture. Each installation uses information as its medium – information which in many cases is hidden or simply not considered in our day to day interactions. Examples

range from implicit associations between ideas and words to more tangible information such as links between Web pages or Closed Captioning (CC) in video feeds.

## 1.2 Relation to Previous Work

Advancements in the use of technology in art in the past twenty years are phenomenal. Computer animation is everywhere, from full-length box office features to animated Flash shorts on the Internet. In gelatin-silver prints, digital darkroom software such as Photoshop and iPhoto moved from the computers of artists to home-marketed bundled deals from Sony and Apple. Illustrator and Painter bring software to 2-D media (rapidograph, charcoal, paint, etc). As the technology gets better, artists become more empowered. However, while useful and ingenious, previously-developed tools are intrinsically limited by their design. They are bound to the space of the media they represent. And while the plug-ins or ‘filters’ are traditionally thought of as a tool for extending the software’s reach, they do not extend beyond its domain. Attempts to go beyond traditional media software are uncommon. They usually require complicated installations, mechanical/physical transformations, and pseudo-immersive environments. As a result, ‘new media’ works are generally static, regardless of how dynamic they may appear. Their actions are either random or hand-tailored. In effect, the system becomes a larger physical instance of a plug-in transformation (blur, sharpen, etc.). Even amongst interactive pieces, the actions tend to be random or tightly scripted.

While a small number of installations have been made in an attempt to reflect media streams [1], we know of no installation or tools that exist which know both about media in the world and media on the computer itself. Many digital libraries hold banks of stock photography and clip art. An information retrieval (IR) system such as Google provides more than just lists of documents, but actually reflects the state of the world, captured as a snapshot of the Internet and what Internet publishers deem popular, interesting, and important. Digital Video Discs (DVDs) provide digitally encoded movies. Even analog television, broadcast over the airwaves, has hidden tracks ignored by most viewers but processed by the embedded computers that play them back. As the digital world becomes more pervasive and computers become more and more invisible, the opportunity exists to build systems which not only leverage all of this newly available information but also act upon it in an artistic manner, creating new experiences for users, and enabling new forms of artistic expression.

## 2. THE IMAGINATION ENVIRONMENT

### 2.1 Watching Television and Videos

The Imagination Environment enters this space as an autonomous emotional amplifier. It watches movies (either on a DVD or TV). While it watches, it searches on-line sources to find images and media clips related to the content of the media being viewed. It presents a selection of the results during its performance. The Environment understands the structure of a scene of video, builds a representation of the scene’s context, and uses that context to find new media. Figure 1 shows the Environment running a performance. The DVD, here *The Godfather*, plays on the center tile as related media is presented in the surrounding tiles. The Environment uses the words and phrases in the dialog to build the context of the scene. It does this by reading (actually decoding) the closed-captioning (CC) information hidden in the video stream, rather than trying to actually listen to the dialog via less reliable speech-to-text technologies [9].

The Environment also knows how its current viewing media is



**Figure 2: A close up of the wall showing the term ‘drink’. Here, a Google image of a do not drink chemical warning is displayed in the upper left corner while an IndexStock image of a child drinking a glass of milk is displayed on the upper right.**

structured. For any closed caption it reads, it identifies the amount of time the caption is displayed, the position on the screen, as well as any hints that are delivered in the stream. Hints are usually the text of audible cues that are provided for the hearing impaired [?]. These cues typically appear in square brackets such as [applause], [whispering], and [gunshots]. In the case of songs, music, and singing, a note graphic, like  $b$ , is placed in the lyrical caption. For DVDs, in addition to the CC information, the Environment uses the DVD’s title and chapter information to identify scenes in the movie while the DVD’s unique identifier, UID, is used to retrieve meta-data from several Web movie repositories, like the title of the movie and its actors.

Once the Environment knows what words are being said and how the media is structured, it uses them to look in several web image repositories to find related pictures. Currently, the Imagination Environment uses three libraries: Google Images [5], Index Stock (a stock photography house) [6], and the Internet Movie Database (IMDB) [7]. Google images are ranked by Internet popularity; the actual image may have nothing to do with the well conceived meaning of the term or phrase. IndexStock, on the other hand, is a hand-picked, human-ordered database, and the images tend to represent canonical meanings of the word. For example, Figure 2 shows how both repositories expand the word ‘drink’. In this case, the stock photo house returns an image of a young girl drinking a glass of milk while Google Images displays a chemical warning prohibiting food or drink. The retrieved image association can be anywhere in the space of the given term. When a movie is talking about an important date, it is not uncommon for the Environment to display pictures of date trees.

Using both repositories together, the Environment expands the space of possible meanings of the word in the video, heightening the visceral appeal of the rhetoric. For example, in the opening scene of *The Godfather*, for example, the undertaker Bonasera is asking the Godfather for vengeance for an injustice which resulted in his daughter being attacked and hospitalized. During his monologue which takes place in the dark mahogany office of the Godfather, he speaks of his beautiful daughter suffering in pain, her jaw wired shut. While he is talking, the images for ‘pain,’ ‘wire,’ and

'beautiful girl' appear around him. Figure 3 shows an example of the images held on the wall during Bonasera's dialog. The visual images within the dark, spoken dialog creates a stronger even more emotionally powerful moment for the Environment's audience.

## 2.2 Presentation, Flow, and JumboShrimp

The Environment can present any type of media: from DVD movies, televised political speeches, to music videos. While the Environment treats each genre the same, each genre's presentation is unique. The subtleties distinguishing each type of media are amplified and made apparent to the viewer. The Environment makes descriptive soliloquies in movies concrete, exposes lexical ambiguities in political speeches, and complements music videos with the imagery in their lyrics. Viewing media that ranges from George Bush's 2003 State of the Union Address, to the Coppola's *The Godfather*, to the music video for Eminem's *Lose Yourself*, The Imagination Environment draws the viewer into an intimate and emotional relationship with both the media and the world of associations and corresponding images it evokes.

Figure 4 shows an example of how the single word 'agreement' can be shown in two contexts. During his 2003 State of the Union Address, George W. Bush refers to Saddam Hussein violating an agreement. At the same time, a Google Image of the Oslo II Interim Agreement is displayed on a neighboring monitor. The Imagination Environment physically makes this juxtaposition by displaying these associations in time with the running media.

It is important to note that not all media moves at the same pace. The speed of a slow dramatic movie monologue does not match that of a live speech or a fast hip-hop video. The Environment balances its rate for presenting images based on the pace of the media and the available presentation space (number of available monitors). Our introductory work in this area creates a model of presentation complementary to the source media. As a result, an effective flow state for the overall installation is automatically achieved.

The actual accounting method varies depending on the structure of the source. For DVD CC information, the Environment looks at how many words in a caption and how many captions are on the screen at once, since each line counts as a caption. It then determines salient words by removing stop words, recognizing characters names, and other such entities. Once it determined the set of terms to display, it looks at the number of available monitors and loads new images over the screens that no longer apply to the current video's context. The rate at which this happens is synchronized with the speed at which the captions are sent in the video stream. To keep the flow state engaging, thresholds are set to keep the images from changing too fast or too slow which prevents the audience from being overwhelmed or becoming bored [3].

The source media for the Imagination environment can be anything text-based. Leveraging its flexibility we created a new instance of the Imagination Environment called JumboShrimp, where the goal is to solely expose the hidden relationships within a body of text itself. JumboShrimp takes as its source any web page, web log (blog), or Internet news feeds via Real Simple Syndication (RSS-XML). In the latter case, salient terms from the news story description are used as the search terms, which are then presented on the wall of monitors. Even though the source is not a constant stream like closed captioning, the flow state is preserved using thresholds tailored to JumboShrimp, allowing the installation to update wall images at a rate which engages its audience. [4]

## 3. ASSOCIATION ENGINE

### 3.1 A Digital Improviser



**Figure 3:** During a monologue from the opening of *The Godfather*, the undertaker talks about his daughter, a beautiful girl with her jaw wired shut suffering in pain in a hospital. The Imagination Environment tiles the images from the dialog and externalizes their relationship to the running movie.



**Figure 4: An example of visually expanding the space of free association found by the Imagination Environment. Here the term ‘agreement,’ from G. W. Bush’s 2003 State of the Union Address, is juxtaposed with a picture of the Oslo II Interim Agreement of 1995, one of the Google Image returns for that term.**

The Association Engine is a hardware/software installation that exposes the intricate web of words that embody language. It externalizes meaningful associations to remind the viewer of connections forgotten and introduce him to new ones. Instead of pictorially expanding links from a term like the Imagination Environment, the Association Engine finds new related terms.

Several embodiments of the Association Engine have been deployed. One such embodiment is based on a warm-up exercise called the pattern game used in improvisational theater. The game is performed by actors standing in a circle. One of the actors says a word to begin the game. The next actor in line does free association from this word. This free association continues around the circle. The goal of this game is to get the actors on the same contextual page before a performance.

To play the pattern game, the Association Engine takes a word from a viewer or the audience and uses that word as a starting point for multi-system free association. A team of machines acts as a group of actors playing the pattern game. Each machine displays a face, which, when synced with voice generation software, becomes an actor in the game. Given a word, a machine searches for connections to other words and ideas using a database mined from Lexical Freenet [8], which indexes multiple types of semantic relationships. This database contains information such as: ‘dream’ is synonymous with ‘ambition,’ and ‘dream’ is part of ‘sleeping.’ The individual machines present these connections to the viewer through both sight and sound, choosing one of the related words as their contribution to the game.

Figure 5 is an artists rendering of the physical installation of the Association Engine. It consists of a combination of flat screen monitors and scrims (transparent cloths used as drops in theaters). The flat screen monitors are placed around the perimeter of the installation. Each screen displays an animated face speaking the words that it contributes to the game. Each face talks and attends to other players by directing its focus on the face that is currently speaking.

Just as the purpose of the pattern game in real-world improvisation is to get performers in the same idea space, the pattern game in the Association Engine creates a common vocabulary reached

by the combined efforts of the individual machines. This bag of words and ideas then becomes the context in which the actual performance takes place. In particular, the performance takes the form of the individual machines doing a One Word Story. In improvisational theater, a One Word Story is performed by a group of actors. One of the actors begins the story by saying a word. In turn, actors add one word to the story at a time.

To create a One Word Story, the Association Engine randomly chooses a story template from a collection of templates. These templates have blank spaces, with specified parts of speech. The Association Engine uses the words chosen during the Pattern Game to fill the spaces. It makes decisions for how to fill the blanks based on parts of speech and semantic relations realized during the pattern game. From a viewer’s perspective, the individual voices trade off to weave the words from the pattern game into a complete narrative.

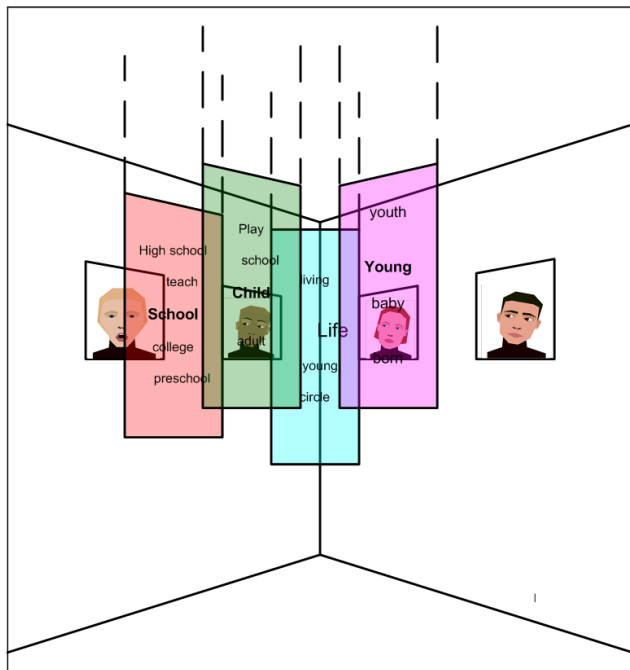
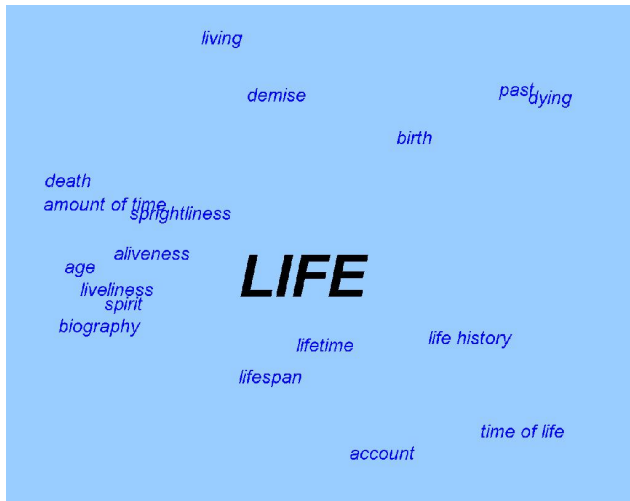
As an example of the Association Engine in action, we will begin with a scenario in which a member of the audience supplies the seed word ‘kitten’, through one of several interaction mechanisms (such as a keyboard, speech recognition engine at the installation, or a cell phone Short Message Service (SMS)). Having received the seed, the faces on the perimeter all turn toward the face on the left. This face says the word ‘kitten’. Following this utterance, the installation displays a variety of related words chosen from its databases of semantic relationships. This collection of related words exposes the kinds of thoughts or contexts evoked by such a single-word utterance in the real-world pattern game. The words evoked by the seed are projected onto the scrims, for example: kitty, puppies, rays, give birth, athwart, young mammal, cat, etc. The next animated face, holding the attention of the other faces, speaks the word ‘puppies’, choosing it from among the many ideas activated by the word “kitten”. ‘Puppies’ emerges from the cloud of projected words, while the rest of the cloud of words disappears. The top image in Figure 5 shows related words expanding in a space. Another cloud appears, made up of words related to puppies: dogs, puppy, kennel, snake, purebred, whelp, pups, collar, cat, breeders, etc. The next animated face speaks the word ‘whelp,’ as ‘whelp’ grows from the group of words. This chain of associations continues as shown in Table 1.

Following the completion of this chain, the virtual players begin a One Word Story. The story is presented in the same manner as the Pattern Game. The individual machines add one word to the story at a time, speaking their addition. As each word is spoken, it is added to the story projected onto the scrims. In this example, the first machine says ‘The,’ the second machine says ‘Trivial,’ the third machine says ‘Gent,’ the fourth machine says ‘A,’ the fifth machine says ‘gent,’ the first machine says ‘once.’ This continues until the complete story shown in Table 1 is read fully by the team of machines.

We believe that this installation provides a strong embodiment for the virtual players while amplifying the notion that they are creating a common vocabulary together. While the players are linked to individual machines, their shared vocabulary becomes externalized in a three-dimensional space of words, ideas, and ultimately a story representing the improvisational experience. The Association Engine, coupled with computer generated faces and scrims as shown in Figure 5, is an installation that opens up the dynamic of team work and performance as a team of autonomous improvisational agents. [4]

### 3.2 Using the Web to quantify word obscurity

In all its embodiments, the Association Engine performs free association using the English language as a corpus. Since the corpus



**Figure 5: Top: The word ‘Life’ is chosen from the set of related expanding terms. Bottom: An artist’s rendering of the Association Engine. The ‘think space’ of associative words are projected on translucent scrims where computer-generated (CG) actors conduct the improvisation.**

### Pattern Game

kitten → puppies → whelp → pup  
 → cup → concavity → impression  
 → chap → gent → spent  
 → idle → laze → loll  
 → banal → trivial

### One Word Story

#### *The Trivial Gent*

A gent once upon a time came forth from his chap in the impression and proclaimed to all the cups that he was a trivial gent, skilled in the use of pups and able to laze all puppies. A kitten asked him, “How can you idle to loll for others, when you are unable to loll your own banal concavity and spent whelp?”

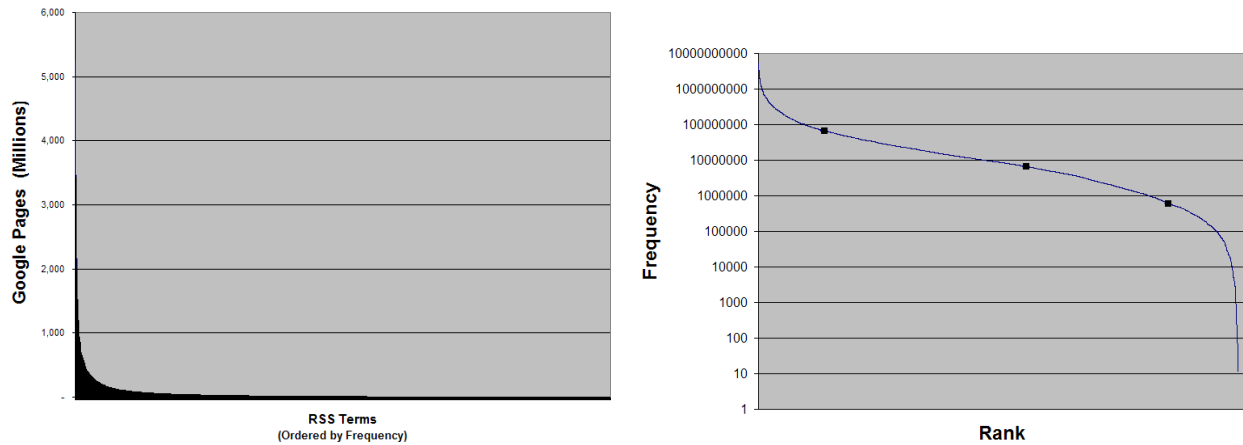
**Table 1: Discovered Word Chain and One Word Story from the Association Engine**

is so large, there are instances where a word chosen for association may be unfamiliar to a general audience. When human actors play the pattern game, they choose words that are recognizable to the other actors. It would be difficult for the other actors to do free association, given a word that they are unfamiliar with. It is effortless for a person to choose words that are not obscure as they are forced to do this in everyday interactions. In conversation, a person must be intelligible, which requires speaking in a vocabulary that can be understood by their audience.

For a machine, determining the obscurity of a word is a nontrivial problem. Our approach is to exploit the Web as an embodiment of the everyday use of human language, in this case, the English language. We hypothesize that the popularity of a word on the Web corresponds to the likelihood that the average audience member is familiar with that word. Using measures of word popularity we drive the pattern game to present a window of cultural understanding, choosing words that are not too obscure and yet not too common.

In order achieve this, we have created boundaries based on the number of search results Google claims to be able to retrieve for a given word. For example, for the query “puppies”, the first page search results from Google, states that it is displaying “Results 1 to 10 of about 2,240,000.” We use the figure supplied by Google as the total number of documents matching a word to determine which words are too common and which are too obscure. To calculate the thresholds by which the Association Engine determines whether or not a word is acceptable, we gathered the document frequency from Google for a sample of over 4500 words from 14 Yahoo! News Real Simple Syndication (RSS) feeds. Graphing the rank of a word against document frequency, we discovered a log normal distribution, similar to a Zipf distribution, Figure 6. Drawing off of properties of a Zipf distribution [10], we calculated the thresholds as one standard deviation away from the average document frequency [2]. With these thresholds in place, the results were encouraging as chosen words were not too common and not too obscure.

A common reaction to this tool is the argument that the Web is a technically biased corpus and will return a higher document frequency for less common, more technical words. For example, a word like ‘orthogonal’ is commonly used in technical reports, academic articles, and other conversations between cyber-geek speakers. If the Web reflects a bias toward technical jargon, we might



**Figure 6: Top: A Zipf distribution of the document frequency of 4500 terms ordered by frequency from the Yahoo! News Real Simple Syndication (RSS) feeds. Bottom: Same graph plotted on a log normal distribution. Terms outside one standard deviation of the mean ( $\mu \pm \sigma$ ) are judged to be too common or too obscure to have any impactful meaning within an installation.**

expect to find a large number of documents using such terms; however, in most cases, Google indexes a relatively small number of documents using these terms (1,460,000 for “orthogonal”), which places their frequency at the lower bound, towards obscurity, of our calculation,  $\approx \mu - \sigma$  and indicates that a bias toward the technical may be small enough to be ignored. However, we realize that this evidence is merely anecdotal and are in the process of conducting a formal study to substantiate the usefulness of document frequency on the Web as a tool for measuring word obscurity.

#### 4. FUTURE DIRECTIONS

The Web plays many roles in our lives. One of the more interesting, yet unexploited, is its role as a storehouse of cultural connections. Search engines, blogs, web portals, and individual web sites are a reflection of our cultural reality. The installations we have described here represent a set of created systems that expose and heighten the connections we use, but rarely see, both in our minds and in the on-line world. By exposing both their results and processes, these systems elevate the mundane, the available, and the purely popular to the level of art. In doing so, the systems themselves are the artistic agents, gathering, sifting, and presenting our own reality back to us as they move through the Web, seeking information.

This new area of *Network Arts* is largely unexplored. At the core of Network Arts are technological advancements in the field of information retrieval, networking and security, social networks, and semantics, but also a cultural understanding of meaning, impact, and artistic portrayal. It is important for the portrayal to be meaningful to the culture it represents and not esoterically complex. Our goal is that in this new form of art and technology, we introduce the machine in art; a role in which the machine is used to expose the world of communication and cultural connections that are linked together and within the grasp of on-line systems. In doing this, a new breed of artists are created, who are able to harness the power of these interconnections to not only create art with the machine but also create artistic agents that themselves are active in the creative process.

#### 5. ACKNOWLEDGMENTS

The authors would like to thank Bruce Gooch and Amy Gooch at Northwestern for their graphics know-how and advice. Ken Perlin for providing us with the improv face work from the Media Research Lab at NYU. Feedback and critiques from several artists on faculty at Northwestern, the Art Institute Chicago, Pensacola Junior College, and several Chicago area artists. Thomas Reichherzer at Indiana University Bloomington for several edits. In addition, continuing thanks to the guiding comments of Larry Birnbaum and other members of the Intelligent Information Laboratory at Northwestern University.

#### 6. REFERENCES

- [1] M. Boehlen and M. Mateas. Office plant #1: Intimate space and contemplative entertainment. *Leonardo: Journal of the International Society for Arts, Sciences, and Technology*, 31(5):345–348, 1998.
- [2] J. Budzik, S. Bradshaw, X. Fu, and K. Hammond. Clustering for opportunistic communication. *WWW 2002*, May 2002.
- [3] M. Csikszentmihalyi. *Flow: The Psychology of Optimal Experience*. Harper & Row, New York, NY, USA, 1990.
- [4] E. Edmonds. Artists augmented by agents (invited speech). Proceedings of the 5th International Conference on Intelligent User Interfaces, January 2000.
- [5] Federal communications commission. <http://www.fcc.gov/cgb/dro/caption.html>.
- [6] Google images. <http://images.google.com>.
- [7] Indexstock images. <http://www.indexstock.com>.
- [8] Internet movie database. <http://www.imdb.com>.
- [9] Lexical freenet. <http://www.lexfn.com>.
- [10] V. Setlur, D. A. Shamma, K. J. Hammond, and S. Sood. Towards a non-linear narrative construction. *Intelligent User Interfaces*, page 329, January 2003.
- [11] G. Zipf. *Human Behavior and the Priciple of Least-effort*. Addison-Wesley, Cambridge, MA, USA, 1949.