# Media fabric — a process-oriented approach to media creation and exchange

## G Davenport, B Barry, A Kelliher and P Nemirovsky

*This paper explores the emergence of a new paradigm we call the 'media fabric' — a semi-intelligent organism consisting of a vast and evolving collection of media artefacts, structures and programs that support our engagement in meaningful real-time dialogues, art making, and social interaction. Situated in and around modern communications networks, the media fabric beckons us into an evolving landscape of creative story potential that is synergistic with our imagination, integral to our everyday life, mindful of itself and our intentions, improvisationally shaped, and provides us with interactions and remaindered artefacts that are evocative and self-reflective.*

## 1.    Introduction — what are media fabrics?

Story creation is an activity of intelligent play, an expressive discovery and exploration into constructing meaning. Storytellers transform real-life observations and fantasies into narratives through acts of selective, inclusion, synthetic emphasis, time distortion and metaphoric encapsulations. The temporal unfolding of the stories present a progressive flow of state changes, vicissitudes of intention shaped within the empowering and framing constraints of their chosen media.

Randomness is a very boring storyteller — good stories are based on a coherent collection and progression of story elements that create larger meaning. Narrative theorists and interaction designers alike have embraced a structural element called the 'story thread'. Each thread is a fragment of a larger story, follows a particular character through some period of time, tracks cause-and-effect chaining of actions over time, or marks the multiple re-emergences of a particular theme. The act of expressing a story thread includes the embedding of a physical and psychological point-of-view, such as that of an involved character, the author, or an 'omniscient' observer. Complex stories are multi-threaded affairs, with many threads juxtaposed as they are interwoven to form a rich tapestry; sometimes, a thread's content is prominently in the foreground, while at other times its content is hidden in the weft and warp of supporting structures.

Over time, the medium of story continually evolves: cinema is a case in point. The celluloid-based genres — fictional movies, cartoons, documentaries, newsreels and advertisements — have adapted and transformed to fit emerging technologies and channels. As an 'old' medium rediscovers itself in digital form, its core becomes collection-based and computationally ready; vast collections of media elements — shots, sounds, still images, and texts — accumulate to be sampled, navigated, sequenced, rendered, repurposed, and exchanged at a moment's notice. Transported over a real-time communications network and experienced as part of everyday life, these media collections have no single author, no beginning, no end. Rather, the collection can be thought of as a 'media fabric' that is ever-present in the environment, a media fabric that is instantaneously available and infinitely reweavable, inviting the participant audience to navigate, create and exchange media elements in ways that are personally significant, satisfying and sociable.

## the future media explorer must move fluidly between consumption and creation

These evolving environments fundamentally challenge traditional delineations between 'author' and 'editor', 'consumer' and 'producer', 'movie' and 'game'. In this media fabric, the future media explorer must move fluidly between consumption and creation. This need challenges the task-oriented digital tools invented in past decades to the breaking point. Both conceptually and practically intimidating to the non-expert, existing tools for audiovisual editing and browsing are designed to be single-purpose and lack the potential for

sociability and playful learning. While editors engage us in the creation of finite artefacts that are completed without leaving any visible or usable trace of the iterative story production process, browsers allow us to navigate almost blindly across story elements without contributing our own structures or traces to the whole. In order to foster the next generation of playful real-time dialogue using rich media, we need to invent a new class of computational instruments, environments and actions that can act as accomplices or provocateurs, that can suggest alternative directions for meaningful play, and ultimately enable the participants to act in a creative tandem with the machine as they navigate this vast and intimate media landscape.

The media fabric is characterised by six critical attributes:

- media construction, exchange, performance and reflection should be **synergistic** and **integral** to our everyday lives,

- it should be **mindful** and **improvisational**,

- it should invite **self-reflection** and **connectedness**.

The reasoning and ramifications of this taxonomy emerge from a high-level historical overview and are concretised as we introduce and discuss three experimental frameworks that tackle problems associated with the media fabric. Each proposes and explores a different paradigm for narrative co-creation, exchange and reflection.

### 1.1    Us++

In Us++, we strive to create a near real-time framework that engages us in a process of reflection and understanding as we record, interpret and share our everyday experiences. In a current experiment, we capitalise on the spontaneous and instant nature of mobile messaging integrating it into the organisational framework of a weblog. This near real-time hybrid communication space promoted playful engagement, enabling a group of friends to stay connected even as one member was in hospital. The potential for rapid, direct cycling of media storytelling and exchange between people will be facilitated as we introduce Mov-its, a custom-designed cell-phone application for gathering, constructing and sharing multimedia messages. By enabling media-rich constructions on a cellular platform, we leverage the 'always with you' attribute of the mobile communication device to support our focus on spontaneous, everyday media-making.

### 1.2    Mindful Documentary

In this project, we seek to build a mindful camera that acts as an intelligent suggestive partner to the human documentary video-maker, thereby extending and reinforcing her art. Using commonsense story models, the camera augments the artist's knowledge and motivates intentional story framing. A combination of human discourse and commonsense story models create a new metadata track that will inform capture and editing decisions. Through this augmentation, our everyday, always-on video device can become a proactive partner in media making, reminding and refreshing the maker's impressions while navigating story potential with a camera and selecting and editing the materials of the media fabric.

### 1.3    Emonic Environment

Inspired by non-idiomatic improvisation, we introduce the option of improvisational interaction within the media fabric to assist the participant in real-time construction and exchange of audiovisual media. In the Emonic Environment, the media fabric is represented as a neural network in which both the content and the structures themselves are subject to the activity of genetic algorithms that control the neural network. Whether used solo, embedded in a sensor-rich space or controlled by multiple users over cell-phones, the Emonic Environment encourages us to engage in the emergence of interesting situations rather than in the planning of a fixed immutable artefact.

## 2.    Background — the transition from media artefact to media fabric

'… living takes place each instant and that instant is always changing. The wisest thing to do is to open one's ears immediately and hear a sound suddenly before one's thinking has a chance to turn it into something logical, abstract or symbolical'.

John Cage (1952)

The artistic, economic, and social history of cinematic media can be broadly subdivided into three major overlapping periods — the movie house, television, and digital. A brief overview of this progression will help to distinguish the unique characteristics of the new media fabric.

The first fifty years of cinema focused exclusively on theatrical release. Cameras, sound recorders, film stock, and projectors were scarce and prohibitively expensive resources controlled by a very small number of producers; thus, story creation was a capital-intensive affair centred on a handful of specially constructed studios, where a factory model of production held sway. The final product — celebrity performances fixed and immutably embedded on to reels of film — were physically transported to appropriately equipped theatrical venues and projected to audiences who had gathered there to simultaneously enjoy a passive, pay-per-view experience.

After the end of World War II, television, which includes the receiving device, channel and content, appeared and proliferated. In broadcast television, the distribution channel was the scarce and expensive resource; precious slices of the terrestrial frequency spectrum were licensed to a few hundred owner/operators of costly broadcast towers. Once established, the channel demanded to be filled with hours of fresh content every day. The technology of television production was considerably less expensive than that of cinema, and hundreds of studios sprang up to make stories for this new medium. More affordable celebrity performances were broadcast live or fixed on to film or videotape for later presentation, and current events such as news and sports could be broadcast in a very timely way. While the movie box office thrived on the 'special night out' social paradigm, television required its audience to stay at home, where different forms of audience interaction and activity — behaviours that would be considered 'rude' in a public theatre — took place around their passive reception of story. Since the television audience 'owned the theatre', the

broadcasters' expenses were recouped by inserting commercial messages into the art.

Even as television was gaining momentum in the 1960s, visionaries such as Marshall McLuhan, John Cage [1], and Claude Levi-Strauss were already forecasting the advent of a more fragmented, personal and creative medium. In *Understanding Media* and *The Medium is the Message*, McLuhan portrays television as a 'hot' medium, simultaneously fragmenting the culture and augmenting individual perception [2, 3]. Around the same time in *La Pensée Sauvage*, Levi-Strauss [4] introduced his now-famous distinction between the bricoleur and the engineer; his characterisation of the bricoleur as one who shuffles and reshuffles elements of a set in order to explore systemic thought has been used by learning theorists such as Seymour Papert [5] and Edith Ackerman [6] to describe children's learning as it is evidenced using constructionist technologies.

In the early 1970s, the VCR and the wireless remote control gave the television audience tremendous new powers of choice, and alternative viewing behaviours began to emerge. Time-shifting, browsing, grazing, and commercial-zapping encouraged new forms of active consumption that exploded the broadcasters' control over the audience's experience of content. At the same time, the proliferation of inexpensive camcorders gave millions of people the power to create their own personal programmes; suddenly, anyone could be a Walter Cronkite, a soap-opera star, or a documentary filmmaker. Video cassettes became a cheap and popular medium for exchange and sharing.

The motivation to create cinematic tools for anyone who had something to say was already present in the designs of early Super8 film systems, the camcorder, and other audiovisual 'consumer formats'. However, this vision was significantly extended into post-production with a promise that every consumer could publish by Apple Computer beginning in the mid 1980s. The success of Apple's campaign to be the computer of choice for visual artists and musicians — professionals and non-professionals alike — generated a profound and lasting effect on the computer industry where the residual impact of products like MacPaint, Hypercard, and System 6 with Quicktime are still overwhelmingly present. While the impressive array of video and sound tools that followed kept the consumer in the driver's seat, these products would not have led us on their own to media fabrics. One more advance was necessary — the ubiquitous, robust network connectivity that took off in the 1990s following the launch of the World Wide Web. Today's array of fixed and wireless digital networks have evolved into an inexpensive, high-speed distribution system available to all; its operation is programmable and personalisable; and, unlike cinema and television, it supports two-way communications and the simultaneous use of as many channels as are needed to interconnect a highly distributed audience.

As was noted earlier, the digital medium is collection-based and computationally ready. When digital imaging and sound technologies were developed in the 1990s, two complementary ideas came to the fore. Firstly, the notion that the consumer can become a teller of tales by actively navigating through a database architecture filled with digital content has struck a responsive chord in formalists and post-modern critics alike. And secondly, by empowering consumers to create and reshape their own content, consumers became authors who would eventually build their own databases as well as contribute to the databases of others [7].

Cultural and technological innovations have resulted in new forms and uses of media. Unfortunately, the very tools that liberated computational expression in the 1990s have now become a limiting factor. Popular composition and editing tools, such as Microsoft Word, I-Movie or WaveLab, have been designed to help people create media artefacts at a very practical level, and browsers are designed without story-like mechanisms that can proactively suggest or affect transitions. The user of a 'nonlinear', random-access video editing system is thrown into a harsh, task-oriented environment where activity is limited to segmenting media and painstakingly positioning the resulting elements into a timeline. There is no concern for the higher-level conceptual play, for browsing through different metadata representations and alternative sequencings, for creating conceptual structures, or for providing the user with subtle or surprising story juxtapositions.

## the digital medium is collection-based and computationally ready

As computer-assisted storytelling systems approach Levi-Strauss' notion of the playful bricoleur, David Snowden points out, 'knowledge is not a "thing", or a system, but an ephemeral, active process of relating' [8]. This suggests that as we move toward a framework for synergistic media construction and exchange, the dialogue with, and appropriation of, story structure becomes a critical gate to the fluidity of the constructional media process.

While it is difficult to predict exactly what will emerge from the confluence of electronic media and communication devices, the 150 million camera-phones, that industry research group Info-trends estimates will be in place by the end of 2004, will undoubtedly result in new applications for media exchange [9]. Based on the rapid rise to popularity of SMS, Firefly, Napster, weblogs and 'flash mobs', one can also predict that new media phenomena will be driven at least in part by a human desire for new modes of social discourse, entertainment and information. With only a small number of camera-phones in circulation today, people can already be observed taking pictures in venues that were formerly picture-proof and dispatching them in an exchange cycle.

As the digital techno age gives way to the digital expressive age, improvisation and media fabrics become an inspiration for story-creating activity. Paul Nemirovsky has defined improvisation as '... a conversation with no end; a conversation that allows us to express our thoughts and feelings, our desire to communicate, think, learn ...' [10]. Nemirovsky's early explorations in emonic improvisational

interaction [11] suggest that creators need to be able to move fluidly between exploratory process and reflection, between temporal structural patterns and particular content, between making, sharing, and exchanging. Broad popular participation in media fabrics requires computational instruments that can act as partner or provocateur, that can suggest and mediate meanings, and that can ultimately free the consumer from having to attend to every detail of a media exchange.

## 3. Media fabric — experimental tools

As our interactions with the media fabric become more frequent and in-the-moment, we are challenged to create novel instruments that invite meaningful play. Salen and Zimmerman provide a descriptive and an evaluative definition for meaningful play: '… meaningful play in a game emerges from the relationship between player action and system outcome; it is the process by which a player takes action within the designed system of a game and the system responds to the action. The meaning of an action in a game resides in the relationship between action and outcome', and, meaningful play '… occurs when the relationship between actions and outcomes in a game are both discernable and integrated into the larger context of the game', (where discernability '… means that a player can perceive the immediate outcome of an action' and integration '… means that the outcome of an action is woven into the game system as a whole') [12].

As we shape our experimental engagement with the media fabric, we seek to build frameworks that can support spontaneous, playful and creative engagement and exchange. This work is premised on the belief that the act of expression as a process is modeled in real-time, and that by bringing capture, exchange, and performance into closer temporal proximity with presentation and reflection, the relationship between expressive action and outcome becomes both discernable and integrated into story potential as it is manifest in the media fabric.

### 3.1 Fun, conversational, collaborative play — Us ++

Storytelling, in its most primitive form, can be described as the simple exchange of experience. In recounting our personal stories we organise and clarify fragments of experienced time by constructing narratives in a dynamic and cyclical process [13]. It is this interplay of agency between the author, actor and the audience that provides us with our sense of self-hood and indeed our narrative identity, both as individuals and as social, cultural beings [14]. In transforming our experiences from raw event to enunciated description, we participate in a formative process of identity construction and refiguration. Bernstein claimed [15] that '… no individual on their own can substantially remake themselves', but, through sharing and divulging our narratives, we engage in a life-long and necessary process whereby we can indeed come to understand ourselves and our actions in a profound and more meaningful way.

Advances in cell-phone technology and increased personal connectivity support the ability of 'ordinary' people to become their own cultural producers as witnessed in the emergence of grassroots, non-corporate Indy-media journalism networks as

well as pranksters using cell-phones to create zany flashmob 'happenings'. In these examples, mobile technology is used to facilitate in the collection and exchange of mediated representations, for purposes of expression, awareness, publicity and fun. SMS messaging points to a new form of mediated storytelling that is mobile, spontaneous, sporadic and immediate. The ephemeral nature of this kind of storytelling, where messages are stored, saved or erased on millions of devices, ultimately means that there may be few remaining discernable traces of these exchanges that are happening very much in the 'now'.

## ultimately free the consumer from having to attend to every detail of the media exchange

Weblogs can offer a viable space for collecting and organising these messages, providing a permanent, chronologically indexed storage location so that these story threads can be further explored and connected, not only by the original targeted recipients, but also by any visitor to the site. In these systems where the roles of author and audience are interchangeable, the rapid cycle of posting and response gives meaning to both actors. The author having posted awaits response; in a moment, audience becomes author; the cycle of relating continues and the process becomes an integral element of the play.

A current on-line experiment, 'Media Fabrics 01', has been designed to exploit the spontaneity and connectivity speed afforded by mobile devices, bringing these together with the organisational structure of a weblog where brief captured experiences can be expanded and further reflected upon. Here, a group of friends are using their cell-phones to correspond and exchange mediated messages with one of their friends who is in hospital undergoing a medical procedure. As the injured party documents their recovery process, the larger group of friends entice the friend towards a future of more normal activities. In this experiment, we strive to create an evolving 'architectured' story process that is built up of many voices and interpretations. On an immediate level there is a desire to connect and extend our physical reach, to provide comfort and companionship to a friend, which is reciprocated by the friend's wish to help remove the alienating distance of the hospital and its attendant medical procedures. The chronological organisation of the weblog, combined with its categorisation mode, makes visible the stages of the recovery progress against a tapestry of the evolving day-to-day interests of the group.

The dialogue nature of this experiment will be extended as we introduce Mov-its, a custom-designed application for the video cell-phone that affords a methodology for gently interrogating lived experiences, where the time-lag between experience, interpretation, representation and display becomes, in some instances, miniscule.

Mov-its, currently in prototype mode, is being implemented in Java for execution on Nokia's Series 60 cell-phones. Creators will be presented with an integrated interface for capturing content, and constructing and exchanging multimedia presentations. Selecting from a variety of default layout templates, participants can populate their template structure with photo, video, text and audio files that have been collected and stored on the telephone (see Fig 1).
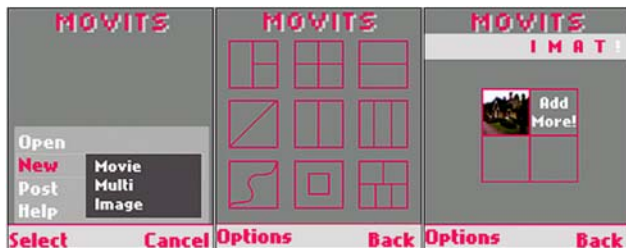


Fig 1    Three screenshots of the Mov-its application. The interfaces shown depict selecting to create a new Mov-it, choosing from the default layout templates, and constructing a Mov-its presentation using stored images and default text slogans.

As the application matures, users will be able to modify existing templates or create their own unique layout structures. The timed display of individual elements in a presentation will be controlled using temporal and transparent variables, allowing the user some creative flexibility within the confines of the template structure. This multimedia presentation —whether complete or invitingly left in a just-begun state — can then be sent using SMS messaging to a friend, a group of people, or published directly from the telephone to a weblog. Recipients of a Mov-it can respond in several ways: they can construct an entirely new presentation to share, they can continue to fill in some elements of the received template before passing the Mov-it on to someone else, or they can retain the received layout structure and temporal settings while adding entirely new content they have collected themselves.

The Mov-its application is deliberately tailored to support a fun game-like process for author and recipient alike. The collected fragments and snippets of activity can be shared immediately, placing the recipient very much in the 'now' of the sender's experience, or can be constructed piecemeal over time where the creator seeks out moments to record that correspond to the idea or theme of their media construction. This approach, which emphasises the separation between media and structure, distinguishes it from currently available commercial 'Moblogs' such as textamerica.com and free on-line media-sharing services like flickr.com that organise uploaded media from mobile telephones using diary or shoebox metaphors, but consider each media element as discrete and not as part of an integrated or interconnecting story. For our experiment, we are currently developing an on-line Mov-its emulator, where visitors to the weblog can also create their own presentations using the content that has been uploaded to the site. In this way, fresh story threads are generated as new connections and juxtapositions are added, and old material is reinterpreted and reconfigured. Through this personal intervention, the material is freed from the original strict chronological ordering of the weblog, and instead offers up potential for discovery of alternative perspectives and interpretations.

The Us++ initiative points to the emergence of a fluid framework for documenting, publishing and revisiting our experiences. As we create, reflect and release our emergent story threads to share with others, we generate story potential for juxtaposition and new meanings in a continuous cycle of play and reflection. The initiative speaks to an assumption and a need: as video cameras and displays become integral elements of our everyday communication device, the overhead cost of commitment to capturing everyday life on video should be dramatically reduced; however, this reduction of overhead requires us to radically redesign solutions for media play, without which video remains locked in the old sequential process that Ralph Rosenblum so colourfully captures in his book *When the Shooting Stops, the Cutting Begins* [16]. As media-rich weblogs provide a framework for rapid publishing that supports chronological and thematic browsing and applications such as Mov-its, small form factor movie-making becomes other than the 'echo of cinema' imagined by Jean Luc Godard; rather, it envisions the media fabric as a fluid, co-constructed and nuanced process, where multiple inputs and perspectives expand and surprise our story-sense in thoughtful and satisfying ways. Freed from the fixed progression of discrete units of activity — shooting followed by editing and then screening — movie-making comes into its own as an integrated, conversational activity allowing the participant to embrace the action that is most appropriate to the moment.

### 3.2    Commonsense construction of cinematic story — Mindful Documentary

For many of us, recording video of our lives, exercises our desire for meaningful narrative play and remembrance. Richard Leacock explains that the documentary film-maker records '… what takes place in the presence of the camera when the camera is turned on' [17]. However, documentary video recording is also challenging, hard work. Insightful documentary making is more than pressing the record button on a camera and walking away. Leacock is suggesting that the film-maker makes a series of decisions about when to record, where to place the camera, and how to frame the shot to best reveal what is happening in the scene according to the film-maker's interest! In short, successful documentary videographers make predictions and track consequences of a broad set of events, situations and characters.

How might this work? At the  beginning of a video capture of a real-time, real-life situation, the videographer establishes a mental model of events and their relationships using the implications of each recorded moment to build complex story organisations. As the recording of a collection of video content progresses, the videographer adjusts the model dynamically to accord with the real world. As models are recalled and revised, connections are made between content elements. Later, after the collection of (often large) amounts of video material, the videographer uses these conceptual connections to make editing decisions during story construction. As video collections grow larger, due to many cameras at a single event or recording, or people recording entire lifetimes, so do the possibilities for narrative construction using those collections.

Can we create a system that can help the videographer capture, organise, navigate and present stories from such large collections of video material? Can such a system use story potential to support maximum story possibility? How will the process of story discovery during documentary videography transition when a camera possesses understanding of the world it is recording and of how stories are constructed?

In the Mindful Documentary initiative, we investigate a process of making documentary videography with a mindful camera that has become an active partner in the process of video capture. Using observations by the videographer as input, the mindful camera compares meta data of shots collected to a vast number of in-camera story representations. Drawing on this wealth of story representations, the mindful camera responds to the videographer's commentary about what is being captured, offering its predictions of what might happen next in the real world. This prediction is processed and returned to the videographer in the form of multiple potential story threads, modest suggestions that speak to what the videographer might capture next. Simultaneously, the videographer's narrative is attached to the video that is being captured. In this model of documentary creation, the mindful camera uses its basic story understanding to assist the filmmaker with three goals:

- capturing unusual details of a documentary subject,

- predicting what might happen next, based on real-world experience,

- generating possible story threads from recorded material.

To imbue a computer with the mental capacities necessary for story comprehension has been a thorny problem for researchers in artificial intelligence. The machine needs resources to reason about the everyday world and methods for identifying or building story structures. Previously, 'expert' systems have been successful at text-based story understanding in extremely limited domains where the engineer has hand-coded representations for all possible story events, details and outcomes in advance. In addition, a very few systems, working in very constrained domains such as cooking shows [18], have used a script-based story representation approach successfully to inform video recording. Recent advances in commonsense reasoning provide a new approach to story understanding — commonsense knowledge resources can be used to support the understanding of broader story domains approximating the real world [19]. This broader understanding is essential for coping with the spontaneous nature of documentary videography.

Inspired by the case-based story-understanding work of Schank and Abelson [20] and Turner [21] as well as in Lehnert's summarisation work [22], Mindful Documentary seeks to use text-based story understanding as an information track that can aid the documentary videographer in real-time story construction. However, unlike an 'expert' system, the mindful camera utilises commonsense resources, including a database of semi-structured stories, to help the videographer

with the task of story collection and synthesis. By polling three commonsense collections, ConceptNet, LifeNet and StoryNet, the camera is able to reason about information that the videographer provides.

## the mindful camera uses ConceptNet to suggest detailed shots to the videographer

These resources can be thought of as systems for reasoning about everyday life. They consist of large collections of commonsense assertions, collected by a large population of non-expert Web users, and mechanisms for reasoning [23—25]. ConceptNet provides a semantic network of commonsense facts that can be used to build context or provide details of a documentary subject. For example, if the videographer is documenting an election, ConceptNet would provide a list of objects used in an election (votes, flyers, commercials, money), places that a person could be found during an election (polls, rallies, fundraisers) and other story ingredients. LifeNet, a probabilistic graphical model, can supply the likelihood of event pairs. If a voter decides on a candidate, LifeNet can generate a list of next possible events and their probabilities. After deciding on a favourite candidate, a voter is much more likely to 'cast a vote' than 'go skydiving'. StoryNet provides a knowledge base of semi-structured stories that can be used for case-based reasoning. An election can be compared to a marathon. Detailed information about these commonsense systems and the reasoning they support is provided by Singh et al [24]. Lieberman et al provide an overview of applications that use ConceptNet [26].

In a current prototype, the mindful camera uses ConceptNet to provide enhanced video metadata and to suggest detailed shots to the videographer. The videographer records a video clip and submits a natural language annotation. The annotation is then used to instantiate a search in ConceptNet for related information, which is presented to the videographer as a future shot suggestion. Metadata for each video clip is automatically expanded, much like in the ARIA photo annotation system [27—28]. The system is designed to be fail-soft; appropriate suggestions can be acted upon, inappropriate ones ignored. A marathon event was recorded using the mindful camera (see Fig 2). This experiment showed that commonsense expansion of metadata might aid system reasoning about events, but the commonsense feedback to the videographer was too unorganised, overwhelming and at a low level of description to aid the documentary filmmaker [29]. We turned to story representations as a more promising resource for feedback but realised that no substantial database of semi-structured stories existed [30]. Research shifted to developing commonsense resources to support Mindful Documentary — specifically OpenMind Experiences and StoryNet — knowledge acquisition projects for collection of semi-structured stories [31]. Currently, these resources are
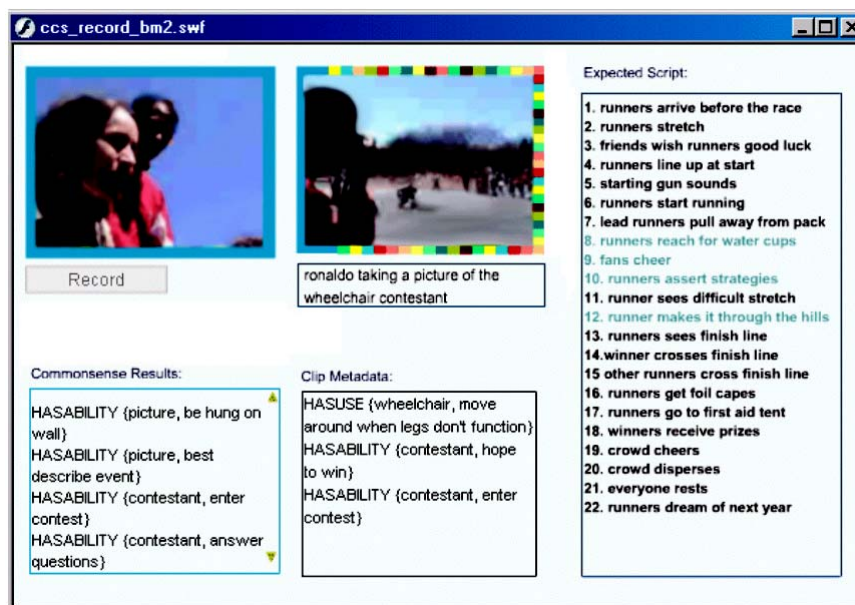
Fig 2    Camera interface for Mindful Documentary showing marathon shoot.

being incorporated into the camera to provide reasoning more appropriate to the task of documentary videography.

The dual meaning of Mindful Documentary is deliberate. The camera mindfully tracks an evolving content collection with an eye on story possibilities. The videographer is actively engaged in story construction with the camera, mindful of story, while participating in everyday life. A critical goal of this work is to help the videographer attend to the construction process while participating in everyday life.

### 3.3    Improvisational approach to constructing and navigating media fabric — Emonic Environment

When we make music, tell stories, or capture images, we frequently do so in the spur of the moment, without carefully considering what each individual note, image, or word might mean. Whether or not we happen to be professional musicians, storytellers, or videographers, we nevertheless are very sensitive to the overall 'flow' of the story — is it rapid, boring, intense, calm? It could be said that we have an intuitive feel for what we consider to be 'good structure' for a given moment yet frequently lack a clear understanding of the structure's details or of the individual steps needed to fill it with words, images, or sounds.

Our creativity often comes in spurts — seeing or hearing something may propel us into a creative mood. Today, however, these creative moods often go unexploited. Why? We hypothesise that a major stumbling block exists in the lack of proper frameworks (and tools) for rapid and creative structural exploration. What can we do to bring about 'every-moment' creativity? We believe that the framework of improvisation presents us with a clear yet powerful answer — design a set of tools that encourages improvisational action, namely:

- allow creation without demanding planning,

- blur the distinction between creation and consumption.

To see why improvisation is an appropriate paradigm, consider what improvisation is.

Improvisation is a conversation with no end — it has no plan nor a fixed set of objectives. Instead, it unfolds in real time, as a combination of action and reaction, producing and consuming. In the domain of human-computer interaction, an improvisational framework allows us to overcome the dichotomy between the two activities, a dichotomy that is manifest in the existing separation between editors and browsers. Participants of an improvisational framework are free to both consume (e.g. by observing input from their environment and from the actions of other participants) and produce (e.g. by responding to the actions of others or initiating their own). Interaction thus becomes a creative activity where even passive actions taken by a participant (e.g. browsing a set of available sounds) can be utilised by others as an input into their creative processes. Furthermore, what improvisers do can be described as discovering latent structures within the improvisers' environment; structures that might go unnoticed in the context of media consumption; in other words, the improviser sees structure where others see only a collection of sounds, images, or words. Perhaps best characterised as a real-time creative action, improvisation involves the creator in a process that is always unique yet not random, that at once encourages co-creation yet benefits solo ventures. These qualities of improvisation make it uniquely suited as a framework for the exploration of the media fabric.

Recently, the topic of improvisation has garnered increasing interest in the research community. One novel direction concerns the design of computational devices that are able to play the role of co-improvisers, thus aiding the improvisers in generation and navigation of both the core artefacts and the structures of a media space. Projects such as Voyager [32], Galapagos [33], ChaOs [34], and Swarm Music [35], explore media spaces in ways that are implicitly or explicitly improvisational. These projects, as well as the Emonic Environment project described below, utilise genetic

algorithms for the creation of media artefacts (the Emonic Environment also utilises the genetic algorithm for structural exploration and path construction).

The Emonic Environment (see Fig 3) is our attempt at creating a multi-participant framework for improvisational action that provides its participants with tools to create, browse, and exchange, in real-time, both the core media and the overarching control structures.

The Emonic Environment presents the participant with a metaphor that reflects the media fabric's interconnected nature — that of a network. Seeing the fabric as a collection of interconnected and evolving media elements and controls allows the participant to focus on creating structural exploration, contributing ideas, and reflecting on the paths suggested by the machine.

Two overlaying neural networks comprise the Emonic Environment architecture:

- a structural network, populated with high-level structural components (called nodes) that define the general activity within the Emonic Environment,

- a perceptual network, populated with lower-level functional components (called emons) that define the processing that can be applied to a given media type.

The behaviour of both the high- and the low-level components is guided by the input from the participants as well as by the built-in genetic algorithms.

## the Emonic Environment presents a metaphor that reflects the media fabric's interconnected nature

The nodes of the structural network are entwined to create a net similar to that of neurons. The nodes each possess a set of properties (activation level, decay rate, propagation threshold, etc). The communication is achieved by the means of generating and passing stimuli between the nodes. Stimuli may originate at any point within the network; they indicate that a given node has become active, i.e. the node's activation level has crossed a boundary, triggering the sending of stimuli
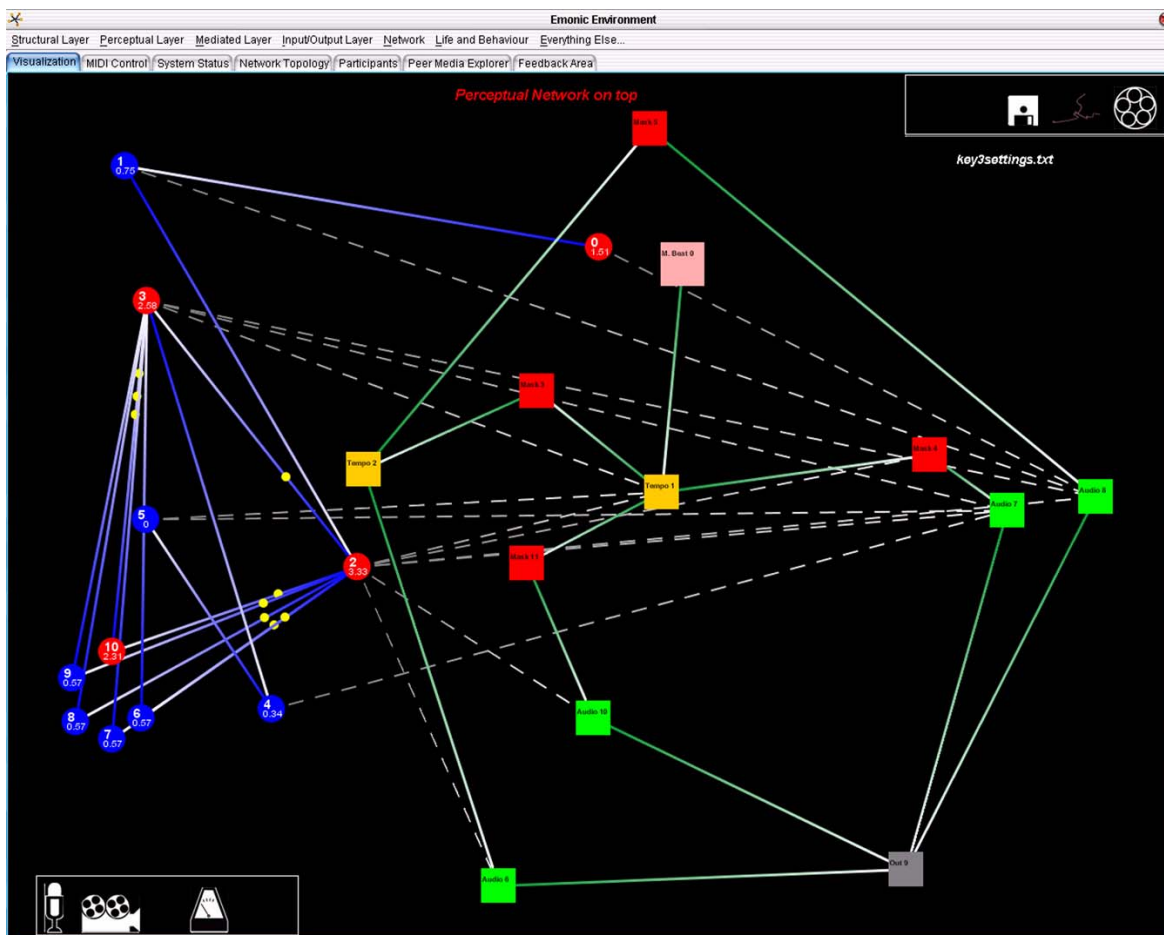


Fig 3    The Emonic Environment running on a desktop computer, with a small portion of the media fabric visible. The nodes (blue and red circles) comprise a neural net guided by genetic algorithms to affect the behaviour of local and remote emons (square elements, colour denoting the type) and their interconnections. The genetic algorithms continuously generate new media fabric states, resulting in a never-ending transition between bricolage structures.

of proportional strength to all the connected nodes. The growth and decay of nodes' activation levels are always due to outside stimulation. Apart from passing stimulation between themselves, these varying levels of activity result in commands being issued to the other, perceptual network, resulting in actual media synthesis, recording, and modification taking place.

In the perceptual network, several emon types exist — for example, one that knows how to play a given audio sample, one that controls when a given action of associated emons (such as playing back an audio sample) takes place (i.e. their behaviour over time), one that knows how to display, rotate, clone, or granulate an image, one that spatialises any given sound, or synthesises one from scratch. All the emons receive input and produce output. That output is then used by other emons as input or released into the media fabric. Four sources of input exist:

- explicit and direct input from a human participant,

- input from other nodes,

- actions taken upon a request from the structural network's emons,

- actions taken upon a request from the currently active genetic algorithm.

The genetic algorithms are capable of running autonomously (i.e. on their own) or being dynamically guided by a participant. The algorithms control the possible mutations of the existing nodes and emons, and define the rules according to which the nodes and the emons can compete and collaborate with each other.

All the elements of the system (that is, the media, the functional elements (emons), and the structural elements (nodes)) are shareable by design. Two modes of sharing are available.

- Traditional mode

    In the first, 'traditional' mode, any subset of the Emonic Environment's media fabric can be selected and sent to another participant currently active within the network. The recipient receives an exact duplicate of what was sent.

- Gift-based mode

    In the other, 'gift-based' mode of sharing, a participant may choose a portion of the available structural network's nodes (see Fig 3), and delegate them to another participant, thus effectively delegating the control of the behaviour of any local emons that receive requests from the nodes in question.

This latter way of sharing allows for a truly shared creation, where the activity happening at any particular workstation running the Emonic Environment may be the result of a real-time collaborative effort by any number of remote participants.

The improvisational foundation of the Emonic Environment's architecture leads to it favouring exploratory processes (interacting with structures) over explicit control of media artefacts, seeking to alleviate the conflict between the immediacy of human creative desire and the inability of our current tools to afford such expression without requiring participants to exhaustively specify all the parameters of what they want to do.

## 4.    Future work

The media fabric manifests itself as media elements and structures that are collection-based, computationally ready and everywhere-accessible. While media fabric encompasses narrative potential of everyone's everyday media, it cannot generate narrative threads on its own. For this, the fabric requires the actions of bricoleurs, humans and machines whose engagement is synergistic and creative, mindful and improvisational, self-reflective and connected.

Collectively the frameworks described above — Us++, Mindful Documentary, and Emonic Environment — help us to envision the future of media use along two critical axes:

- the social context of our engagement with media and its structures,

- method and interface for search and organisation of media.

From an action-oriented perspective, the media fabric promotes engagement in process rather than production as a finite artefact. In our journey within the fabric, we are influenced by the tools we possess — a telephone, a camera, smart clothing. Traditionally single purpose devices, these objects will increasingly be designed to function as both media collector and controller.

> the genetic algorithms are capable of running autonomously or being dynamically guided by a participant

In our future work, we attempt to integrate these devices with ever more meaningful organisation and search mechanisms implemented in ways that capitalise on our improvisational passage within the fabric. Currently each framework emphasises a single mechanism. The diary format of Us++ provides the human maker with a strong mnemonic structure which can be used to review and reflect upon individual and social memory. Mindful Documentary incorporates commonsense knowledge of everyday life in its semantic approach to navigating narrative potential. The Emonic Environment explores ways in which the human and machine come together to play the role of bricoleur — offering genetic algorithms which can lead us to discover unique media configurations. These approaches are complementary and

when combined will potentially draw strength from each other, helping to make the media fabric synergistic with human desire for engagement, discovery and understanding.

Many participants who access the media fabric have had little or no media experience. As we progress, we strive to give the fabric a presence that generates trust as well as spontaneous action so that it can inspire all participants to explore, construct, and share media without knowing (or worrying) where they are going or precisely what they will experience.

## 5.    Conclusions

We began this paper by making the point that storytelling transforms real-life experience into a communicative experience using the empowering and framing constraints of a chosen medium. Storytellers create story threads as they engage in a process we can construe as meaningful play. Story creation, sharing and exchange are driving activities for communications networks.

A brief history traces the impact of miniaturisation of production technology, computation and the move towards large collections of media fragments as digital media blurs the traditional roles of 'author' and 'editor', 'consumer' and 'producer', 'movie' and 'game'. We claim that these collections will increasingly present themselves as media fabrics, landscapes that engage the media bricoleur in meaning, making activities that are synergistic, integral to daily life, mindful and improvisational, and that invite self-reflection and connectedness.

The media fabric we propose requires new media instruments that would allow us to engage in meaningful play as we navigate, create, and exchange media throughout our daily lives. A definition offered by Salen and Zimmerman [12] frames our discussion of three parallel research inquiries into new instruments and processes for playful activity within the media fabric. US++ concretises the idea that stories can be generated incrementally, as if in conversation with an audience [36]. As many people in a community contribute to this dialogue, the story becomes rich, expansive and multi-threaded. With Mov-its, US++ leverages the intimacy and ubiquitous adoption of the cell-phone, which increasingly incorporates potential for rich media creation and viewing, to allow for a more improvisational framework for the co-creation of story threads.

The mindful camera addresses a daunting problem that has faced videographers ever since video cameras appeared on the consumer market: how can we best discover, capture and construct a meaningful cinematic story about the real world? By drawing on common-sense representations of the storied world, mindful camera seeks a creative partnership between the computational camera and the videographer. This approach supports active story inquiry during capture and produces a collection of story representations to accompany the video content; this process helps the videographer search, select and construct within the media fabric.

Finally, work on the Emonic Environment highlights the role improvisation can play in reframing our everyday creative interactions with the media fabrics. A neural network representation provides access to the structural core of the media fabric; genetic algorithms generate navigational paths; human players explore the paths, design and exchange media and structures, and introduce corrective feedback to the actions of the machine as they choose.

Each of these instruments strives to position real-time or near real-time creative action in a way that results in meaningful, playful and frequently sociable engagement. While each explores how computation can help people navigate, create and exchange media in a more immediate way than is possible today, each is also clearly framed by the idea that storymaking is a  process involving both humans and machines in a synergistic and evolving engagement.

## References

1   Cage J: 'Silence; lectures and writings', Cambridge, MA, The MIT Press (1966).

2   McLuhan M: 'Understanding Media: The Extensions of Man', New York, McGraw Hill (1964).

3   McLuhan M: 'The Medium is the Message', New York, Random House (1967).

4   Levi-Strauss C: 'La Pensée Sauvage', London, Weidenfeld and Nicolson (1966).

5   Papert S and Turkle S: 'Epistemological pluralism and the revaluation of the concrete', Journal of Mathematical Behavior, 11, No1, pp 3—33  (March 1992).

6   Ackermann E: 'Piaget's constructivism, Papert's constructionism: what's the difference?', Constructivism: uses and perspectives in education, Vols 1 and 2, Conference Proceedings, Geneva, Research Center in Education/Cahier 8/September 01, pp 85—94 (2001).

7   Davenport G: 'New Orleans in Transition, 1983—1987: The Interactive Delivery of a Cinematic Case Study', Proceedings of the International Congress for Design and Planning Theory, Education Group Conference, Boston, MA (1987).

8   Snowden D: 'Introduction to the cynefin centre for organisational complexity', (2002) — http://www-1.ibm.com/services/

9   Sharma Dinesh C: 'Global sales of camera phones shoot up', CNET News.com  (March 2004).

10  Nemirovsky P and Watson R: 'Genetic improvisation model: a framework for real-time performance environments', Proceedings of the EvoGP 2003, Gloucester, UK (2003).

11  Nemirovsky P: 'Aesthetic forms of expression as information delivery units', MIT Masters Thesis (1999).

12  Salen K and Zimmerman E: 'Rules of play: game design fundamentals', Cambridge MA, The MIT Press (2004).

13  Ricoeur P: 'Time and Narrative', 3 Vols, Chicago, University of Chicago Press (1984).

14  Kearney R: 'On Stories', New York, Routledge (2002).

15  Bernstein J M: 'Self-knowledge as Praxis', in Nash C (Ed): 'Narrative in Culture: The Uses of Storytelling in the Sciences, Philosophy and Literature', New York, Routledge (1990).

16  Rosenblum  R and Karen R: 'When the Shooting Stops the Cutting Begins', New York, Da Capo (1979).

17  Leacock R: 'Ethnographic observation and the super-8 millimeter camera', in Hockings P (Ed): 'Principles of visual anthropology', New York, Mouton de Gruyter (1977).

18  Pinhanez C and Bobick A: 'Approximate world models: incorporating qualitative and linguistic information into vision systems',  Proc of AAAI'96, Portland, Oregon, pp 1116—1123, (August 1996).

19 McCarthy J, Minsky M, Sloman A, Gong L, Lau T, Morgenstern L, Mueller E, Riecken D, Singh M and Singh P: 'An architecture of diversity for commonsense reasoning', IBM Systems Journal, 41, No 3, pp 530—539 (2002).

20 Schank R and Abelson R: 'Scripts, plans, goals, and understanding', Hillsdale, NJ, Lawrence Erlbaum Associates (1977).

21 Turner S: 'The Creative Process: A Computer Model of Storytelling and Creativity', Hillsdale, NJ, Lawrence Erlbaum Associates (1994).

22 Lehnert W: 'Plot Units: A Narrative Summarisation Strategy', in Lehnert W and Ringle M (Eds): 'Strategies for natural language processing', pp 375—414, Hillsdale, NJ, Lawrence Erlbaum Associates (1982).

23 Singh P, Lin T, Mueller E, Lim G, Perkins T and Zhu W: 'Open mind common sense: knowledge acquisition from the general public', Proceedings of the First International Conference on Ontologies, Databases, and Applications of Semantics for Large Scale Information Systems, Irvine, CA (2002).

24 Singh P, Barry B and Liu H: 'Teaching machines about everyday life', BT Technol J, 22, No 4, pp ...... (October 2004).

25 Singh P: 'The Public Acquisition of Commonsense Knowledge', Proceedings of the AAAI Spring Symposium: Acquiring (and Using) Linguistic (and World) Knowledge for Information Access, Palo Alto, CA (2002).

26 Lieberman H, Liu H, Singh P and Barry B: 'Beating some common sense into interactive applications', To appear in AI Magazine (2004).

27 Lieberman H, Rosenzweig E and Singh P: 'Aria: An agent for annotating and retrieving images', IEEE Computer, pp 57—61 (2001).

28 Liu H: 'Semantic understanding and commonsense reasoning in an adaptive photo agent', Masters Thesis, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA (2002).

29 Barry B and Davenport G: 'Documenting life: videography and commonsense', Proceedings of IEEE International Conference on Multimedia, New York, Institute of Electrical and Electronics Engineers, Baltimore, Maryland (2003).

30 Mueller E T: 'A database and lexicon of scripts for ThoughtTreasure', CogPrints ID cog00000555 (1999) — http://cogprints.soton.ac.uk/

31 Singh P and Barry B: 'Collecting commonsense experiences', Proceedings of the Second International Conference on Knowledge Capture (K-CAP 2003), Sanibel Island, FL (2003).

32 Lewis G: 'Interacting with latter-day musical automata". Contemporary Music Review, 18, No 3, pp 99—112 (1999).

33 Sims K: 'Galapagos', Information online at: http://www.genarts.com/galapagos/ (1993).

34 Miranda E R: 'On the Origins and Evolution of Music in Virtual Worlds', in Bentley P and Corne D: 'Creative evolutionary systems', San Francisco, CA, Academic Press (2002).

35 Blackwell T M and Bentley P J: 'Improvised music with swarms', Congress on Evolutionary Computation, CEC-2002 (2002).

36 Davenport G and Murtaugh M: 'Automatist storyteller systems and the shifting sands of story', IBM Systems Journal, 36, No 3, pp 446—456 (1997).

Glorianna Davenport is head of the Interactive Cinema group at the Media Lab. Trained as a documentary filmmaker, she has achieved international recognition for her work in new media forms.

Her research explores fundamental issues related to the collaborative co-construction of digital media experiences, where the task of narration is split among authors, consumers, and computer mediators. Her recent work focuses on the creation of customisable, personalisable storyteller systems that dynamically serve and adapt to a widely dispersed society of audiences. She received a BA in English from Mt Holyoke College, and an MA in fine arts from Hunter College in New York City.



Barbara Barry is a PhD student in the Media Lab's Interactive Cinema group. Her research explores how video cameras equipped with artificial intelligence can have an impact on how we document our lives.

Currently, she is developing commonsense knowledge bases and reasoning techniques to further necessary story understanding by computers to support documentary video construction in a world of always-on, always-ready cameras. She earned a BFA from the Massachusetts College of Art and an MS in media arts and sciences from MIT.



Aisling Kelliher is a PhD candidate in the Media Lab's Interactive Cinema group. Her work focuses on personal cinematic story construction and exchange. She is developing tools and technologies to support media fluency and promote the emergence and sustenance of active, reflective communities. These systems will provide a space for representing our daily experiences and depicting our memories as stories that define who we are both as individual and social beings. Before coming to the Media Lab, she taught multimedia programming, digital video, and vision/music technology at Trinity College, Dublin. She holds a BA in communication studies from Dublin City University, an MSc in multimedia systems from Trinity College, Dublin and an MS in media arts and sciences from MIT.



Paul Nemirovsky is a PhD student in the Media Lab's Interactive Cinema group. He is interested in making improvisation the dominant framework for creating, searching, and interacting within the digital domain. As an undergraduate, he was interested in music and design, as well as computer and cognitive sciences. He has performed and composed extensively and internationally and worked as a lead programmer in a number of high-tech companies in Israel.

While at MIT, he worked on wearables, media search engines, and multiple projects for entertainment and learning.