



InterPARES 2 Project

International Research on Permanent Authentic Records in Electronic Systems

Title: Case Study 09(3) Final Report:
Digital Moving Images – Commercial Film
Studio

Status: Final (public)

Version: 1.1

Submission Date: August 2004

Release Date: September 2007

Author: The InterPARES 2 Project

Writer(s): James M. Turner, Université de Montréal

Marta Braun, Ryerson University

Mary Ide, WGBH Boston

Randal Luckow, Turner Broadcasting System

Michael Murhph, Ryerson University

Isabella Oreice, Associazione Nazionale Archivistica Italiana

Andrew Rodger, National Archives of Canada

Project Unit: Focus 1

URL: http://www.interpares.org/display_file.cfm?doc=ip2_cs09-3_final_report.pdf

Table of Contents

Introduction.....	1
A. Overview of the Case Study	1
B. Statement of Methodology.....	2
C. Addressing the 23 Core Research Questions.....	2
D. Addressing the Domain Research Questions.....	14
Domain 1 Questions.....	14
Domain 2 Questions.....	16
Domain 3 Questions.....	18
 Appendix 1: Workflow Diagram	 20

Introduction

This is the report of the first part of the investigation conducted within the framework of case study 09 of Working Group 3.1 of the InterPARES 2 Project. Over time, the case study became known as “the case study that just grew and grew,” since it began with a single investigation and over time included the study of three additional institutions. The research team discussed at length various possibilities of modeling the processes of pre-production, production and post-production in filmmaking and of identifying the documents (called “digital assets” or “assets” in the industry) produced at each stage, and eventually opted to study commercial production, on the grounds that this activity would offer the most complete and most generic model. A large commercial film studio agreed to be our partner for this investigation, on condition of anonymity. This report represents the fruit of our work on this first investigation.

In this report, we provide our answers to the 23 core research questions and the Domain 1, 2 and 3 research questions to which all InterPARES 2 researchers are required to respond in their case studies. These answers represent the data collected on site at the commercial studio, as subsequently revised and edited by members of our research team and the chair of our working group.

Reports of additional activities of our research team, conducted with our other partners, the National Film Board of Canada, Altair4 di Roma, and WGBH Boston, are forthcoming. The suggestions for including additional institutions in our study, offered by various members of the team, easily won the enthusiastic support of the rest of our team because these offered the possibility of gaining a much broader understanding of the processes involved in various types of filmmaking activity as well as in multimedia production, than we would have gotten from the study as we originally conceived it.

We thank Yvette Hackett for her watchcare, for her advice as chair of working group 3.1 within InterPARES 2, and for her many, many helpful suggestions for improving this document.

A. Overview of the Case Study

The anonymous large commercial (theatrical) film studio employs around 2000 artists, marketing, publicity, legal, administrative and other support staff necessary in the creation of a modern CG animated feature. The studio is well recognized and has won many awards, including an Academy Award.

This case study focuses on the workflow necessary to create a theatrically released CG animated feature. From early pre-visualization, through story, and onto final moving images, this case study follows the digital and physical instances of image assets from concept to screen. Sound and non-moving image material was not evaluated as we considered this outside the focus of study.

B. Statement of Methodology

The case study was completed by the Archivist of the anonymous commercial film studio in a series of interviews with various department heads, tracking and describing the assets as they are manipulated and passed along down the workflow from pre-visualization to final film out. The Archivist created a document to visually represent the movement of digital assets through the creative workflow of the studio. This document allowed the team to easily understand the creative processes in a step by step manner, facilitating communication within the team.

Further work and answers to the case study questions was completed by the case study team, quizzing the Archivist, with the additional participation of Yvette Hackett.

C. Addressing the 23 Core Research Questions

1. What activities of the creator have you investigated?

Our study investigates the sequence of steps taken in creating the animation in a “cg” animated film (i.e., an animated film made using computer graphics). Starting in the development stage, it covers the visual aspects of an animated production from conception to final product, including some of the non-digital activities which continue to be part of the process. The investigation excludes all sound and the editorial processes, concentrating on the artistic aspects, including painting and drawing.

The production process can create many “traditional” types of digital records such as memos, spreadsheets, and so on, but these are excluded from our case study. The digital entities being studied are the artwork which is modified and merged until it becomes “the work,” which is subsequently distributed digitally (e.g., in DVD format) or on film (e.g., in a theatrical release). Additional assets are created for marketing purposes. The ratio of digital to non-digital (i.e., analogue) entities will undoubtedly vary depending on the production company and the type of animation being produced (for example, a short, a feature-length movie, an insert).

2. Which of these activities generate the digital entities that are the objects of your case study?

The question is complex. All the activities generate entities, some of which begin as analogue and are later converted to digital. The main activities under study are visual development, character design, and prop design. Generally these activities begin with analogue entities created using pen and pencil with paper. These are scanned to produce a digital entity, although there can be exceptions. These digital entities constitute the material sent around to the various players for a variety of treatments. In the visual development stage, there is much manipulation of the digital entities.

All computer files generated are of interest to our study. However, there is much mixing of analogue and digital elements in the processes, and they cannot be isolated from one another. For example, a design might be done on cardboard, and then scanned to a .TIFF file so the

production staff can view it. But the processed file is then printed to prop sheets which receive an approval signature from the director. Thus there are physical assets at the beginning and the end of this process, with a computer file in between. The fact that the physical manifestation bears the signature makes this version the document of record which is archived. The digital files may have a continued operational value to the creator, but without the signature they do not have value as records.

In the creation process some of the physical assets have a digital file behind them, and some do not. This probably varies from one studio to another. All of these pieces could be created as digital files using computer programs. However, the practice is to make 2D pieces the primary assets because artists are more used to these; drawing with pencil and paper is much faster for producing drawings. In addition, in meetings, or when showing a number of drawings to those who need to see them, manipulating physical objects is much more efficient than manipulating digital objects. Thus the digital objects are used mainly for adding effects, refining lighting, and so on, and the physical objects remain the primary archived assets.

3. For what purpose(s) are the digital entities you have examined created?

All the assets under study are specific to the department in question at the time the modifications to them are made. Each manipulation of the digital files is done in order to advance the production. Each player in the workflow process receives the digital file from the last person, completes the required artwork or other manipulation of the file, and then sends it on to the next person. Scene painting, background creation, lighting effects are all separate products that are compiled to form the final image.

The major process functions include:

- visual development to get the look of the film;
- story treatments to define the story;
- story panels to define the timing and action (rough movement);
- animatics, created using Avid software, to further refine the timing and make the movement fluid;
- character design to achieve the approved look of each character, after which it is prototyped in the modeling department, finalized and released for the animators to use. The animator then has an object to manipulate (this type of animation is analogous to stop-motion puppet animation);
- layout, which defines the equivalent of camera movement in live-action film and establishes the placement of animated characters and props in each scene. This process also defines the painting requirements; and
- lighting, which is exceedingly important in CG animation; the work is done in a minimum of three steps, or passes.

In the process of animation, there are not a great number of digital entities created, but rather digital entities that evolve. This involves constant modification to existing files, each iteration being composited and flattened to reduce the file size. Compositing and flattening have to do with taking the new artwork that is created as an additional layer on top of the previous iteration, and merging them together to form a single new layer. It is rarely necessary to revert to previous

versions. However, previous iterations are saved for a time in order to permit this if it becomes necessary. Eventually they are overwritten.

File saves are carried out in the following way: files that are modified are overwritten. Thus not every file is saved every day, only the ones that have been modified are overwritten. There is a cyclic arrangement. Eventually inactive files are sent to off-line storage. Since important hardware and software changes have usually occurred since the artwork was created, in the environment we studied it is deemed more economical to re-create artwork if it is to be reused than to engage in a process of migrating or otherwise upgrading it in case it will be reused. This approach limits the amount of time, energy and money spent on long-term preservation of digital entities and avoids the need to implement preservation strategies that respond to problems of hardware and software obsolescence.

4. What form do these digital entities take? (e.g., e-mail, CAD, database)

The software used includes Maya, Render Man, Word, and Photoshop. E-mail is used very little in these processes, and no email files are kept after ninety days.

Files created by the Maya software include the extensions .ma and .md. These files have to do with wireframe and other initial versions of artwork, and .ms files (also produced using Maya), are those which contain the rendered surface artwork. Rigging is contained in the .ma or .md files, but it is not clear which of these it is exactly.

Original artwork is usually scanned into TIFF files, and animatics, produced by the Avid software, are created in the form of .mov files. After all the artwork has been created, modified, lit, and rendered, the final output for the finished film is in TIFF format. This last stage is accomplished using Shake software.

The Render Man base software was purchased early on, and then modified to suit the needs of the studio. Many in-house tools have been created, including software to open any digital image file format. Thus the software has become proprietary, although the studio may still be paying for licensing the source code.

As many assets are created for marketing purposes as for purposes of animation of the film. Since production usually has a 4-year timeline, material produced in the first year can be stored offsite for lack of space onsite.

4a. What are the key formal elements, attributes, and behavior (if any) of the digital entities?

Word is used for word processing. Maya is used for character modeling, rigging, animation, and surfacing. Render Man used for rendering, and for lighting as well. Shake is used for compositing.

Storyboard Panels are scanned as JPEG files in order to provide the Avid software with a digital file to manipulate. The Avid software is used to work out the timing and as reference to create the animatic .mov files. These are used to show timing and fluid

movement (in the form of .mov files). Photoshop is used for visual development and production artwork (such as background paintings). Shake is used for compositing and for exporting final TIFF files for film output.

4b. What are the digital components of which they consist and their specifications?

In the animation process, all digital components are geared to creating the final output TIFF files. The process can be viewed in a number of ways. Some aspects need to be tagged for archiving, such as the initial matte paintings (however, these are not wholly digital components). Each piece of a word file is archived, such as included Excel tables, included images, and so on.

Analogue material is kept largely because of its commercial value. Individual animation cels or other artwork can be sold in galleries or at auction, but digital files have no value or only ephemeral value for such purposes. Ironically, digital files can take on value if they are touched by famous people, for example a celebrity may have written on the label of removable media even though he/she never used the computer with which the file was created.

At one point in production, the image is combined with the soundtrack, which consists of other digital components. However, the editorial process is outside the scope of our study.

4c. What is the relationship between the intellectual aspects and the technical components?

The intellectual aspects are the artwork objects, and the technical components are the tools used to manipulate these objects. The content of a Word document is the intellectual aspect, and the technical component is the software used to open the document again and thus provide access to the intellectual information. The intellectual existence of the artwork is in the image file, and the technical aspect of it resides in the software used to manipulate the image.

4d. How are the digital entities identified (e.g., is there a [persistent] unique identifier)?

Strict naming conventions are used to identify the digital entities, and all those having a role in manipulating the file are required to adhere to these conventions. Among other elements, the name of the file contains information on the sequence, the scene, the name of the object, as well as numerical information to identify the version.

The sequence of information in the file name is as follows:

- /studio/title/sequence/scene/object/version.extension

Interpretation of this information is as follows: “Studio” refers to the name of the studio that owns the artwork, since occasionally artwork is outsourced to another studio or a subsidiary. “Title” refers to the working title of the film being produced. “Sequence” and “Scene” refer respectively to these parts of the film (in the parlance of the studio we studied, “scene” is the equivalent of “shot”). “Object” refers to the particular piece of

artwork in hand. Finally, a version number is added to identify the precise iteration of the file.

Sometimes in pre-visualization or at the story stage, there is an abbreviation to include information such as the sequence date and the name of the artist.

There has been some attempt to develop a consistent taxonomy. Specific terms to describe each object in development are selected in the brainstorming stage by the production team. Thus there is agreement by committee on the naming conventions to be used for each production. These, however, do not extend from one production to another.

4e. In the organization of the digital entities, what kind of aggregation levels exist, if any?

The answer to this question resides in the path recorded in the name for each digital object. Information is aggregated in the file name, including studio, title, sequence, and so on.

4f. What determines the way in which the digital entities are organized?

The workflow needed to produce the film determines all the processes. Objects are passed from one worker to the next for manipulation and treatment, following precise sequences, in order to achieve the final output image files in TIFF format.

5. How are those digital entities created?

The initial digital entities are created by scanning artwork so the Avid software can create a story reel. This is replaced wholly by the animatic and the animatic is replaced by animation. Computer animation is very much a process of manipulating computer files and rendering them each time to flatten the layers of art work and reduce the file size. All processes are a function of the workflow. Each time a file is rendered, some new digital entity is created. Although it is saved in addition to the previous version of the same digital object so there can be some kind of backup in case the need to return to the previous version arises, in practice the previous versions do not survive very long.

5a. What is the nature of the system(s) with which they are created? (e.g., functionality, software, hardware, peripherals, etc.)

The animatics are created in the form of .mb Maya files and published as .mov files. In addition, libraries of objects are built and stored using the same Maya software.

Almost all hardware consists of Hewlett Packard computers using Red Hat Linux as the operating system. All the software (Maya, Render Man, etc.) runs in this environment.

Computers are linked in a large server network. Apache software is used for security, and there may be some Novell functions for sign-ins.

Peripherals include printers, and plotters, used mainly to create presentation boards for visitors, artists, meetings and for investors.

Hardware in the studio is subject to frequent change. The arguments of newly hired employees for moving production to another computing environment are considered, and if the arguments are convincing enough, the system is changed.

5b. Does the system manage the complete range of digital entities created in the identified activity or activities for the organization (or part of it) in which they operate?

Yes.

6. From what precise process (es) or procedure(s), or part thereof, do the digital entities result?

Processes and procedures that produce digital entities include: Animation system, Animation character integration, Visual development, and Shot create. The workflow diagram in Appendix 1 offers a graphic representation of this information.

7. To what other digital or non-digital entities are they connected in either a conceptual or a technical way? Is such connection documented or captured?

The story panels which are used for the pitch (presenting the idea for a film to decision-makers in order to gain their approval to make the film) and which are in analogue format have numbers to identify individual images. When the images are scanned into the Avid computer, these numbers appear and are used as references to the corresponding images on paper. Scripts are clearly named and the path to these files on the computer's hard disk appears as a header on the cover page of the script, along with the last revision date.

Entities created in the visual development stage are almost invariably objects in analogue format. No connection is made between the object in hand and the scanned, photocopied, or other interim versions of the object. Such objects can be considered variant proxies of the digital object but no connection is made between them in production.

Nor are entities created in the character design stage connected, because they are imported for reference when a model is created, but are discarded thereafter. All these elements are analogue entities.

8. What are the documentary and technological processes or procedures that the creator follows to identify, retrieve, and access the digital entities?

These processes are the naming protocol and path. The studio is the creator. The answer to this question resides wholly in the naming conventions used by the studio. It has taken the studio many years to come to understand the importance of these conventions. Only now is the use of them understood.

9. Are those processes and procedures documented? How? In what form?

A training department exists; however, staff work from handouts. Procedures are constantly in flux. No central procedures manual exists. When a new production begins, the new staff creates a structure loosely based on previous production experience they have had, rather than from any sort of procedure manual. Although there is no manual, the strict naming conventions alleviate problems this would normally create, since everyone involved is made to understand the importance of using these conventions correctly. Thus training and information about passing objects from one person to the next is done by word of mouth. The use of word of mouth techniques is the documentation, which can be considered a kind of oral tradition.

10. What measures does the creator take to ensure the quality, reliability and authenticity of the digital entities and their documentation?

The version number gives authenticity to the digital entities, and assigning this correctly is assured through training in the use of the naming conventions. The quality and reliability of digital entities are assured if the digital entity is usable in conducting the next stage of work. If the document responds to the software first by opening then by responding to the software's commands, it is considered reliable and of appropriate quality.

If the day's work of each person is recorded in the system, it then becomes available for the next person's use. The primary trust is in the security of the system. At the system level, the backups made daily constitute the assurance of reliability.

11. Does the creator think that the authenticity of his digital entities is assured, and if so, why?

Yes, the creator is certain the authenticity of the digital entities is assured, because if they are not authentic they will not register correctly on top of previous artwork, nor will the compositing operation be successful. This is the digital equivalent of animation cels registering on pegs on an animation stand and of colors registering correctly inside the boundaries of line drawings when the two elements are not created in the same layer, for example when a number of cels are stacked, each bearing different elements, to create the image.

12. How does the creator use the digital entities under examination?

The creator uses the digital entities in a number of ways, some involving new creation. In the course of production, all activities are geared to producing the final output TIFF images. Apart from the initial development stage, almost all animation activities are carried out using the digital entities.

However, activities do not end with this. Legal aspects are involved in recording the day and date, as well as the author of the work. Legal aspects constitute the primary use of the digital entities after the film is released, and these are largely centered on ownership.

In addition, there are a number of marketing and promotional uses of both still and animated images from the film. They are the basis for producing merchandising materials. Other marketing

and promotional uses include re-use of digital entities for interviews with animators, party events, awards, and so on, such as value-added material on DVD versions of the studio's films.

13. How are changes to the digital entities made and recorded?

The image of making a quilt is used internally to describe the process of constructing the film. The production starts off as a roll of blank film, and as pieces are constructed, they are “sewn in” to the blank film, in the space where they need to occur according to the story. Each piece is made independently of the others and “sewn together and placed in the quilt” or Story Reel.

Elements are versioned as they are built. Throughout the quilting process digital entities are completely overwritten by the next version. At each level of the processes of storyboards, animatics, animation, compositing, to the final TIFF images, there is final artwork. However, since that artwork is not the final product for the film, it is considered merely a stepping stone. When final approval is achieved, the versions are eliminated, unless a special request to archive them is made.

The Technology Department's activities support the work during production, and then they are gone. Since there is no system for digitally archiving these materials, nothing can be retrieved from storage unless the file name and path are known, as well as an approximate date of production. The process is somewhat akin to detective work. Because of the many layers of work in each digital entity, it is highly doubtful that a specific element could be retrieved from a precise version, from a particular date.

Once the production is completed, digital entities are stored on backup tapes in some kind of order. However, since there is no records management program, selecting material to save is like choosing material floating by on a moving river. All that is saved is simply whatever the archivist manages to select.

During production, activities take place on a tight schedule, so if it becomes necessary to restore digital entities from backup tapes, it is a question of days, not weeks. The backup system is based on catastrophe scenarios, not on records management, so that the system just happens to have material from the previous days.

14. Do external users have access to the digital entities in question? If so, how, and what kind of uses do they make of the entities?

Yes, marketing, publicity and consumer products departments, in liaison with an office called Animation Production Support, have access to the digital entities. This office is internal to the company, but external to the production. Here artwork, textual material and stills can be had to help in the production of publicity (such as advertisements and one-sheet posters) for the film, and for merchandise related to the film. The artwork can be anything from script synopsis and visual development, to frames of the film still in progress, or frames that have received final approval. Access is given to specific files through an approval process.

Outside companies make the trailers for the films produced by the studio. They receive material from the studio but must return it afterwards. The material can be shots from the film, or from

any stage of production back to the animatics that are produced early on. These transactions take place within a secure environment with the vendor, and the work is controlled by security contracts. Any violations such as allowing the material to be used for any other purposes than those specified in the contract ensure that the outside contractor will never get another job from the studio.

The Apache software manages approval of requests to the directory level, so that there are internal controls on the permissions available. External users do not have such access. Security is set up at the directory or folder level, and governs the production database used to check these assets in and out. The database can also manage notes, which can be passed among employees along with the assets. The directories have an information structure and a tracking mechanism to provide control over the movement of the assets the database contains.

There is an older database/workflow application, which was previously used for tracking 2D animation work. It remains administratively useful for CGI films in a slightly different incarnation.

When the Shot Create command is given, a structure is created but it has no content. It is a placeholder where items can be put. In a sense, this is the equivalent of a field which does not yet contain a value. Shot Create places an empty directory structure on the production servers awaiting asset creation and delivery.

15. Are there specific job competencies (or responsibilities) with respect to the creation, maintenance, and/or use of the digital entities? If yes, what are they?

Animators have expertise in specific areas, but skill sets are such that there is much overlap among them. All this work involves close contact with the animation technology, which is used to manage the digital entities, back them up, assigning them an address in the system, and keeping them responsive to hardware and software structures. Much of the work is governed by union rules, which delineate the responsibilities of each animator, as well as the functions, and the compensation received for each type of work.

16. Are the access rights (to objects and/or systems) connected to the job competence of the responsible person? If yes, what are they?

Yes, each user of the system has specific permissions, depending on the responsibilities they have for producing part of the film. Security is set up at the directory or folder level, and governs the database used to check assets in or out.

17. Among its digital entities, which ones does the creator consider to be records and why?

Only files saved and sent to the archives are considered records. All other materials are considered ephemeral, and exist solely to advance the work of the production. Digital entities are sent to the archives for legal and marketing reasons, because some further use is anticipated. There is no program for records management, but archiving takes place when there is some specific reason for it.

18. Does the creator keep the digital entities that are currently being examined? That is, are these digital entities part of a record keeping system? If so, what are its features?

Although there is not a record-keeping system in the company at large, there is one in the archive. The system manages the following functions: import, tracking and output. It was purchased as a turnkey DAM system, tailored to the studio's needs, and re-branded. It is used as a research tool by the studio.

All assets from feature animation that need to be used by the marketing department, the consumer products department, and outside vendors are put in DAM system and administered by the archive. Features include ingest, transform, the ability to e-mail and FTP assets securely, and to aggregate sets of assets. The system can be considered a repository and a digital asset management system. However, the studio does not intentionally hire trained information professionals such as librarians and archivists to manage the system, so the personnel that are hired fend for themselves and invent the information management processes, often rather clumsily.

The final output artwork in TIFF form is stored in the system, along with the appropriate metadata including the sequence, the scene (the studio's terminology for the shot), the frame, as well as the title of the production, the division that delivered the assets to the archive. There is no plan for rich metadata such as subject indexing because in the context of the company and for the purposes for which the information system is managed, this is not necessary because all assets are current. Physical artwork is collected with the understanding the artwork will be given as gifts or sold.

Despite the system's weaknesses, some standards are used by the Archive, such as the Dublin Core and the Categories for the Description of Works of Art (CDWA). The tracking system is based on the CDWA and is considered adequate for the studio's purposes. The Categories have been installed as they appear in the standard, and the archivist has created a set of rules about how certain fields of the database need to be populated using these authority records. However, much explaining and persuasion need to take place for users to adopt these practices, since the archivist does not have the authority to enforce these desirable practices.

18a. Do the recordkeeping system(s) (or processes) routinely capture all digital entities within the scope of the activity it covers?

No, none of this activity is routine. All capture to the system is done manually. The archivist needs to learn of digital entities that should be archived, then gather them, and finally ingest them into the system.

18b. From what applications do the recordkeeping system(s) inherit or capture the digital entities and the related metadata (e.g., e-mail, tracking systems, workflow systems, office systems, databases, etc.)?

Another database, built with FileMaker Pro, is used for tracking physical pieces of artwork that are not digital.

18c. Are the digital entities organized in a way that reflects the creation processes? What is the schema, if any, for organizing the digital entities?

Yes, the organization of the digital entities does reflect the creation process. Starting from a series of animatics created for the film, the digital entities are grouped in series such as lighting keys, digital X-sheets, and so on. Because of this arrangement, users can know where the pieces fit in the hierarchy and into the work flow.

Identifying terminology exists, based on the creator's name and place in the company structure. The various departments create assets, and the archivist maintains the list of identifying terminology. However, since the list is highly confidential, it is not available.

18d. Does the recordkeeping system provide ready access to all relevant digital entities and related metadata?

Yes, access is maintained for all relevant digital entities and their metadata. Everything in the system that can be opened can be downloaded.

18e. Does the recordkeeping system document all actions/ transactions that take place in the system re: the digital entities? If so, what are the metadata captured?

No, for the moment only the check in and check out transactions are documented. The transactions that modify a record's metadata are not documented at present.

19. How does the creator maintain its digital entities through technological change?

The creator does not maintain its digital entities through technological change. Long-term preservation is not a concern. The culture of the studio is such that until a situation arises that creates an actual problem, the question will not be addressed. Digital assets have a steep drop-off in usability after the first tier media window is exhausted.

19a. What preservation strategies and/or methods are implemented and how?

There are no preservation strategies or methods because it is felt that none are needed. The studio is only ten years old and the culture of the organization is such that there is no thought of preservation or migration. If an older digital object is needed but cannot be accessed because of technological change in the meantime, new artwork is created or some other piece is used instead. Legal reasons are the motivation for archiving anything at all.

19b. Are these strategies or methods determined by the type of digital entities (in a technical sense) or by other criteria? If the latter, what criteria?

There are no such strategies or methods.

20. To what extent do policies, procedures, and standards currently control records creation, maintenance, preservation and use in the context of the creator’s activity? Do these policies, procedures, and standards need to be modified or augmented?

While there are some policies that try to control the workflow of record creation and quality, once the file has moved on there is no such thing as preservation, except for a global backup. Assets that make it into the archive are saved; however, preservation of these assets is temporary. In addition, there is no migration activity to ensure that these assets can be accessed in future. The policies and procedures need to be modified, but there is no institutional motivation for carrying this out.

21. What legal, moral (e.g., control over artistic expression) or ethical obligations, concerns or issues exist regarding the creation, maintenance, preservation and use of the records in the context of the creator’s activity?

Legal obligations have to do with making contracts, proper crediting of the work in each production, and respecting union guidelines. As is the case in general with the Hollywood studios, there are few moral or ethical obligations, concerns or issues. Legal obligations are the only ones to which any attention is paid.

22. What descriptive or other metadata schema or standards are currently being used in the creation, maintenance, use and preservation of the recordkeeping system or environment being studied?

There are no standards for creation of the assets in the workflow. However, the archivist has introduced standards for description and indexing which cover those assets that make it to the archive. These include the Categories for the Description of Works of Art (CDWA), the Dublin Core (DC), the Thesaurus for Graphic Materials I: Subject Terms (TGM I), the Thesaurus for Graphic Materials II: Genre and Physical Characteristics Terms (TGM II). The Anglo-American Cataloguing Rules are used to describe scripts, manuscripts, partial notes, and such. Some tracking information about other documentation is recorded using the Turabian Style Guide and The Chicago Manual of Style.

23. What is the source of these descriptive or other metadata schema or standards (institutional convention, professional body, international standard, individual practice, etc.)?

Institutional convention governs practice during the workflow stage for any particular production. A snapshot of the entire directory structure for each production is kept, but users trying to access materials from even recent productions have been unsuccessful because of hardware and software changes that occurred in the meantime.

Material that is archived is done so using the tools listed in the answer to question 22, so professional bodies and international standards govern these activities.

D. Addressing the Domain Research Questions

Domain 1 Questions

1.1 (a) What types of documents are traditionally made or received and set aside (that is, created) in the course of artistic, scientific, and governmental activities that are expected to be carried out on-line?

Static artwork.

(b) For what purposes?

Visual design and story definition.

(c) What types of electronic documents are currently being created to accomplish those same activities?

Mostly Adobe Photoshop files and scanned tiffs of physical artwork.

(d) Have the purposes for which these documents are created changed?

No.

1.2 (a) What are the nature and the characteristics of the traditional process of document creation in each activity?

Physical artwork created on paper or canvas.

(b) Have they been altered by the use of digital technology and, if yes, how?

Somewhat. Physical artwork can be scanned and manipulated to create variations.

1.3 (a) What are the formal elements and attributes of the documents generated by these processes in both a traditional and a digital environment?

Not applicable.

(b) What is the function of each element and the significance of each attribute?

Not applicable.

(c) Specifically, what is the manifestation of authorship in the records of each activity and its implications for the exercise of intellectual property rights and the attribution of responsibilities?

Artwork is signed and dated. All assets created during the course of production are owned wholly by the studio, not the creator.

1.4 (a) Does the definition of a record adopted by InterPARES 1 apply to all or part of the documents generated by these processes?

No.

(b) If yes, given the different manifestations of the record's nature in such documents, how do we recognize and demonstrate the necessary components that the definition identifies?

(c) If not, is it possible to change the definition maintaining theoretical consistency in the identification of documents as records across the spectrum of human activities?

Unknown.

(d) In other words, should we be looking at other factors that make of a document a record than those that diplomatics and archival science have considered so far?

I think my education in archival science may be at odds with InterPARES's interpretation.

1.5 As government and businesses deliver services electronically and enter into transactions based on more dynamic web-based presentations and exchanges of information, are they neglecting to capture adequate documentary evidence of the occurrence of these transactions?

Not applicable.

1.6 Is the move to more dynamic and open-ended exchanges of information blurring the responsibilities and altering the legal liabilities of the participants in electronic transactions?

Yes. Legal inquiries have requested the Archive to print digital documents for review.

1.7 (a) How do record creators traditionally determine the retention of their records and implement this determination in the context of each activity?

In this instance, record creators have no influence over retention.

(b) How do record retention decisions and practices differ for individual and institutional creators?

Record retention is the prerogative of the institution.

(c) How has the use of digital technology affected their decisions and practices?

It hasn't.

Domain 2 Questions

2.1 (a) *What does record reliability mean in the context of artistic, scientific and government activities?*

Usability.

(b) *To what extent can the electronic records created in the course of each type of activity be considered reliable and why?*

Fully reliable, if they open and can be used.

(c) *What requirements on their form and controls on their creation would make us presume that they are reliable?*

None.

2.2 (a) *What does record accuracy mean in the context of each activity?*

Accuracy is based on image registration and final version.

(b) *To what extent can the electronic records created in the course of each type of activity be considered accurate and why?*

Fully, if image registration occurs and/or the versioning is final.

(c) *What controls on their creation would make us presume that these records are accurate?*

The only control is based on following versioning notation correctly, if not, accuracy is based solely on registration.

2.3 (a) *What does authenticity mean in the context of each activity?*

No meaning.

(b) *To what extent is the definition of record authenticity adopted by InterPARES 1 relevant to the records resulting from each type of activity and from the use of increasingly complex digital technology?*

None.

2.4 (a) *On what basis can the records created in the course of each activity be presumed authentic?*

This is a closed process, internal to the company and behind firewalls. If an object, asset or record is in the workflow, it is authentic because it exists in this realm.

(b) How, in the absence of such presumption, can their authenticity be verified?

Not applicable.

2.5 (a) How is the authenticity of these records affected by their transmission across space and time?

Across space authenticity is based on file naming conventions, across time authenticity is based on inclusion in a dated backup file held on a DLT.

(b) What controls on the process of transmission would ensure that these records will continue to be recognized as authentic?

File naming conventions.

2.6 Are the conceptual requirements for reliability and authenticity developed by the UBC-MAS project [Duranti and MacNeil, 1999] and InterPARES 1 for administrative and legal records generated within databases and document management systems applicable to the records studied by InterPARES 2?

Unknown.

2.7 (a) Do the participants in electronic transactions have shared access to reliable and accurate information about the terms and effects of the transactions?

No.

(b) What would constitute reliable and accurate records of transactions in current electronic service delivery initiatives?

Unknown.

2.8 What would be the consequence of issuing guidelines for record creation on the nature of the records of each activity?

The culture of the studio frowns on any restrictions to the creative process.

2.9 How can cultural differences, freedom of expression, freedom of inquiry, and right to privacy be reflected in those guidelines?

Unknown.

2.10 What technological and intellectual tools would assist creators to generate records that can be authentically preserved over time?

Cataloguing tools and meaningful dispositions to the Archive repository based on a current Collection Policy.

2.11 What legal or moral obligations exist regarding the creation, use and preservation of the records under investigation?

Some records are preserved in response to legal needs, usually based on crediting and other union rules. There are no moral obligations.

Domain 3 Questions

3.1 What types of entities does the diplomatic analysis identify in this case study? (i.e. records, publications, data, etc.).

No diplomatic analysis has been done for these records.

3.1a Are the entities reliable? If not, why not?

Reliability is based on the ability of Animation Technology to provide access to files that can be used during the course of artistic endeavour. If the entities (files) can be opened by the artists in the workflow, they are reliable. If the files cannot be opened they are not reliable. Reliability is based on usability. This is rarely an issue of concern because all entities are current.

3.1b Are the entities accurate? If not, why not?

Yes. Because the entities are visual, they must register properly and act upon other entities correctly. Accuracy is visual and errors in registration or action (or interaction with other entities) can be seen, and become a problem to be resolved in the software. All entities must be accurate. All versions must be accurate for the same reason, because everything must render properly.

3.2 For what purposes are the entities to be preserved?

Preservation is an activity that is carried out during the course of production, to serve production for re-use and disaster planning. Once the sequence has been rendered to film, preservation is no longer necessary for production. Also, once the entire film has been output, preservation and disaster planning is no longer necessary. Few digital assets are held separately from the DLT back-ups, snap-shots and series of output tiffs. These few assets make to the Archive for Marketing and Publicity usage. Traditional assets continue to be collected, mostly for legal purposes (Intellectual rights, ownership, authorship, etc.).

3.3 To what degree can the entities be presumed to be authentic, and why?

While there is no cross check to verify the authenticity of the entities collected on the DLT's, they are considered authentic simply because this is a closed environment and the pipeline directs these entities to be written to the DLT's at the end of their active life-cycle (after film out).

3.4 Given our knowledge of current preservation methods, can the entities be preserved?

Possibly. But this is not a concern for the organization. Hardware and software obsolescence has already rendered many of the files in the back-ups useless. All of the DLT's related to the motion capture approach to *Shrek*, are unreadable, and were unreadable within three years of creation. Certainly there was also an issue of an application of a proprietary file type that was used early on as a security tool, but because it was undocumented and staff turnover was complete, there was no way to 'crack' the code. The application of a standard to all file types preserved and a consistent migration scheme would help ensure the entities are preserved. Alternately, since the final output tiff is considered the most valuable entity, and that this is a ubiquitous file type, they are easily preserved—and this is being done purposefully—all other preservation is a by-product of the pipeline. And most likely, because of the harshness of the work environment, back-up post the completion of production is simply departments and individuals avoiding the responsibility of deletion or record destruction.

3.5 If yes, what elements and components need to be preserved?

Final output tiffs are considered primary, and are preserved to DLT, or may remain for an unspecified amount of time offline.

3.6 Which preservation method might most usefully be applied, and what are its strengths and weaknesses?

As in question 3.5.

3.7 What metadata is required to support appraisal and preservation? If metadata is missing, where should it come from and how should it be made manifest?

There is no way to find files in the DLT's other than by browsing, if you don't know the exact file name you are looking to find. The structure of the film is based on the hierarchical arrangement of Sequence/Scene/Frame, so if you have a sequence list and knowledge of the film, you should be able to get to the proper frame, if you can figure out which DLT it resides on.

3.8 Are there any policy constraints that would affect the preservation of the entities?

There are no policies in place on the organizational level regarding preservation of entities. Traditional assets are continually mined to be given as 'gifts' by the organization. As mentioned above, the pipeline workflow designates certain snapshots be made of the directory structures during the course of production as well as writing the output tiffs to DLT. But there are no migration policies or reliability checks in place post completion of the production.

Appendix 1: Workflow Diagram

See next page.

