Interview with Ben Fino-Radin

> Ben is a New York based media archaeologist and conservator of born-digital and computer-based works of contemporary art. At MoMA he is managing the design and implementation of the Digital Repository for Museum Collections, as well as advising on the collection and conservation of media based artworks and design objects. At Rhizome at the New Museum, he leads the preservation and curation of the ArtBase, one of the oldest and most comprehensive collections of born-digital works of art. He is near completion of an MS in Library and Information Science and MFA in digital arts at Pratt Institute, with a BFA in media art from Alfred University.

April 26, 2013

Interviewers: Crystal Sanchez and James Smith

Please describe your experience working in the field of time-based art and its preservation.

My undergraduate degree was in Fine Arts with a concentration in New Media – real time video synthesis, sound, sculpture. At Alfred University, the New Media department was founded by professors who were very much part of the Western NY video scene in the 1970s that revolved around the Experimental Television Center, lots of real-time video synthesis, and stuff like that. That informed my initial experience in this field; I began as an artist engaged with technology as material, craft, and tool. After three years of making and showing work and playing in some bands in Providence, Rhode Island, I wound up working in the archives of the Hasbro Toy Company. There, I fell in love with archives, which led me to discover some of the interdisciplinary graduate programs in New York. In 2010 I moved to New York to attend Pratt Institute for a Masters in Library Science and an MFA in Digital Arts.

During my first year of graduate school, I began a research fellowship at Rhizome—a New Museum affiliated non-profit institution dedicated to the creation, presentation, preservation, and critique of emerging artistic practices that engage technology. Although I had only just begun my graduate degrees I devoted most of my time and energy to the research I was conducting at Rhizome. Unsurprisingly, I found the work (and the freedom) to be far more challenging and rewarding than my graduate studies. After about two years at Rhizome, Glenn Wharton, MoMA's first (and now, former) Media Conservator approached me about managing the development of the digital repository for MoMA's digital collections – a project that he had led for roughly three years. Today I am at MoMA three days of the week, managing the development of the Digital Repository for Museum Collections (DRMC), and working on various other media conservation projects, including the collection and preservation of the video games that have recently entered the

collection, as well as a very exciting case-study on emulation in the museum setting. I am currently still at Rhizome on a part-time basis, finishing up some projects there, but as of September 13th, 2013 will be leaving my post and devoting myself full time at MoMA.

Please tell us about Rhizome and what your role is there.

A huge way that Rhizome carries out its mission is through its archive. It truly is the only institution doing the sort of strategic collection development and historic preservation— going deep into the history of art and technology to find the forgotten histories and untold stories, preserving not only a collection of born-digital artworks, but preserving historian's ability to observe how computers entered the practice of contemporary artists. We also are actively and aggressively collecting contemporary work, and developing meaningful and lasting relationships with artists, so that we may work with them to understand how we may stabilize their projects as close to the point of creation as possible, and in an unobtrusive manner. For instance, when we collected *I'm here and there .com,* by artist Jonas Lund, we worked with the artist to collect the server-side components of the work, duplicated his database, and he suggested the idea of modifying his browser extension that sends the browser data so that it now communicates with both his database for the original URL, but also the database hosted on Rhizome's infrastructure that powers our mirror of the site.

Rhizome is also unique in the legal manner through which works enter our collection – we operate on a basis of a non-exclusive license that grants us rights for distributing the work, and documentation and metadata about the work freely, but also explicitly preserves the artists' right to do so. We don't want to own the intellectual rights and we want to preserve the artist's ability to sell and distribute their work as they see fit. In essence what we hold is the preservation master. The one sense in which this model does not apply to the collection is with more unique, one-of-a-kind materials, such as a vintage computer containing important historic materials. We have collected these historic machines from various collectives (such as The Thing BBS, or ArtNet BBS) so that we recover and preserve the materials contained on them, but we are at the point where we have to content with whether or not we want to be in the business of retaining and collecting the physical artifacts that are these machines themselves.

Do you have different preservation strategies for the various types of works in your collection?

Yes. As the very contemporary (as opposed to 20-30 years past creation) work is completely born-digital and not stored on any sort of obsolescent or at-risk storage media, we do not have to conduct any sort of digital forensics, data recovery, or emulation. Ninety percent of the work that we do is with living artists making work today, the collection process all happens online and very seldom do we have to sneaker-net files by shipping a

hard drive or flash drive. The footprint of web based materials is so small. Rafael Rozendaal donated his entire body of work of the last thirteen years, which was well under a gigabyte. Furthermore, many artists we collect are already sharing their code openly and publicly through sites like GitHub or Gitorious. No one is easier to collect than an artist that is firmly grounded in open-source practices, as this means generally speaking that their materials are easily at hand, and already documented for the public.

For the new acquisitions, have you standardized a workflow for things that you ask the artists for or the questions that you ask them?

While we always like to document our naturally occurring correspondence with the artist about the technical makeup of a work, and conduct a conservation interview whenever possible, I would not say that at Rhizome we have any sort of formalized set of questions that we bring to every scenario. I am of the opinion that successful collection practices must rely on the conservators' domain and material expertise. It should be trivial for a conservator of digital materials or media based artworks to extrapolate what they need to ask of the artist after spending some time researching the work. I would caution against relying on formalized questionnaires.

Do you think that you could formalize polices, and at what level- at the level of classes of works?

I think the only level that blanket formalization can occur is at a high level that touches on legal, ethical, and very basic preservation issues. For instance – at Rhizome, when we agree to acquire a piece, the first thing to occur is sending the artist a release form or a deed of gift. With something so subtle and variable as a work of art, I don't think this is possible to the point of having SIP requirements or SIP classes [submission information packages] and then automatically checking those to see if – here's this database dump and now we can ingest this to the repository, we don't do it at this level and I don't think this is possible for complex software based works. That entire OAIS paradigm was designed for a much different context – one for interoperable repositories of space data. It can work for archives that may be conducting a mass digitization project, and know precisely what types of materials they need in a SIP. It is silly to apply this to something that is as variable as art. I think that the one way this can apply to our practice is in the documentation of works. For example, at MoMA, we have recently formalized what sort of documentation is created when a media based work is exhibited. This includes assembling a package of floor plans, documentation of decisions that were made, a video of a walkthrough of the exhibition, and other such materials. In this sort of scenario I could absolutely see us checking for SIP completeness, but again, I don't think that this mentality applies to actual artwork materials. There are some general basics, such as if it is compiled software, get the source

code – if the work references a database, get a dump of the database, but there are much more specific collection and conservation procedures that are very contextual and specific to the work at hand.

Can you talk about the technical infrastructure at Rhizome and MoMA?

At Rhizome, we operate on a quite a shoestring budget. I would not call our archive a repository. I would call it a very carefully maintained and controlled file system with limited access. All of our public materials are hosted on an Apache server in an Amazon EC2 instance. This is constantly backed up locally at the New Museum on virtualized storage infrastructure managed by the New Museum's IT staff. We also have a portion of this storage environment dedicated to dark archives of materials that are not yet public, or may never be. For the most part, everything that Rhizome collects is not solely for the purpose of preservation, but primarily access. For instance, when we collect a web-based work we immediately toss it on our server so that it is live. It is only recently that we have been working with older, historic materials such as The Thing BBS, where we are imaging entire computers and other devices that may contain personal or sensitive information. There is a bit of analysis and curation that needs to occur before these materials are thrown out into the wild. In a sense, Rhizome is growing up and having to contend with quite standard archival concerns with regard to curation, redaction, and privacy.

To compliment this somewhat basic infrastructure, I have recently completed the design of a Collective Access based collections management system for Rhizome. This new system is providing Rhizome with major upgrades for collections oriented metadata, such as enforcement of CDWA-lite for the description of artworks, more robust metadata about artists and people related to works in the collection, but also for the first time at Rhizome will be combining the massive archive of Rhizome events and exhibitions. The real win with this new system though is that we are being a bit unconventional, combining our collections management and preservation metadata. If an artwork is dependent on a particular operating system, an object is created in the database for that operating system, and it is related to the artwork through an ontology that describes the nature of this relationship. We are essentially creating a linked network in our collection management system of technical requirements. For example, a record for Netscape 3 would have a relationship to the operating systems it is compatible with and then those operating systems have relationships to the emulators that they can run in, or vintage hardware that we have that they can run on. This means that when you are cataloging a work from 1996 and the artist requires that it be seen in Netscape 3, you can establish that relationship and then immediately see what you need to display the work. That is something that is valuable to researchers, so that is why we decided to include it in our collection management system. It is a very human-driven system, but I think it will prove to be quite effective.

Our environment and context at MoMA is of course entirely different. At MoMA I manage what is more truly a repository, and what is essentially akin to art storage for digital objects. The Digital Repository for Museum Collections (DRMC) holds all born-digital and digitized materials in MoMA's collection. This ranges from software based works of contemporary art, to digitized film, and from video games, to digitized video art. We are in the process of building an Archivematica and AtoM based system that communicates with TMS, and facilitates the collection, preservation, and long term access of these materials. We currently hold over 40TB of digital collections, and this is growing at a rather rapid pace. Our infrastructure is a virtualized storage pool running on EMC hardware mirrored between two geographically redundant locations, with backups made to LTO.

This system when complete will be much more automated than the aforementioned Collective Access system, and all description is inherently tied to the Archival Information Package (AIP). To put it in very simple terms – when we receive new materials from an artist, and a conservator has decided that we have everything that we need and these materials are ready for long-term storage, they process them in Archivematica. We are embedding certain metadata in these AIPs that allow the repository management application to automatically correlate the AIP with whatever artwork in the collection it belongs to. This management application is not simply for access, but becomes a place where these AIPs are enriched with further metadata. We can similarly describe a dependency on a particular operating system, but we are doing this at a very low level, describing that this Artwork AIP has this relationship to this Supporting Technology AIP. Of course because all of this description is inherently tied to the original digital objects, which also possess verbose characterization metadata, we envision this management application becoming an incredibly powerful digital conservation management tool for many other institutions. I'm pleased to say that the results of MoMA's work will be free and open source, and become standard modules of Archivematica and AtoM.

Could you talk about the taxonomies, or categories Rhizome uses? Is it something that people can build on?

I think that in the late 90's and early 2000's there was a lot of discussion about taxonomies in the preservation and classification of time based works of art – and this was a conversation that Rhizome was very much leading in. However, if you look at how this has panned out it is clear that much of this interest in taxonomies and "folksonomies" at the time was a bit of an unhealthy combination of pseudo information science, and web 2.0. The thought of archive patrons or artists tagging materials was talked about as this radical

idea, as though it was solving some kind of problem of classification or preservation. I don't think this to be the case one bit.

In my opinion, ontologies are far more important for the needs of describing the complexities of a work of contemporary art that is reliant on technology. If you look at the Collective Access profile that we've developed at Rhizome, it is entirely predicated on exposing and describing the relationships between artworks and technologies as discrete intellectual objects. These relationship types in our ontology between artworks and technologies are things such as: Minimum System Requirement, Known Compatibility, Used as Material, Contained in Artwork, Preferred for Display, etc. These are not all of the artwork/technology relationship types in our ontology, and I imagine that as Rhizome continues to use its new collections management system that these values may evolve and grow on an as-needed basis. To take it a step further, we also have the ability to when needed, describe relationships between technologies themselves. For instance, should we describe a work as having compatibility with a specific version of a specific web browser, we could describe specifically which operating systems that web browser is compatible with. This depth of cataloging is not needed in all cases of course, but this is all to say that for the needs of conservation of time based works of contemporary art – to me the need is not for taxonomies or categories, but rather ontologies. Even if you look at standards as verbose as PREMIS, there are only very limited suggestions for controlled ontological values. Despite all of the attention that linked data is seeing in the information science field these days, I am not aware of the existence of or initiatives for the formalization of such ontologies as I am describing. The interesting thing about Rhizome's new setup is that all of this metadata is inherently also part of the access system. This information that is needed for preservation also in the end becomes an incredibly powerful tool for researchers to discover and map the evolution of artist's relationships to materials and technologies over the decades.

I appreciate that this preservation and descriptive metadata are so dependent on each other that they are blurred.

Well, there is of course preservation metadata that you wouldn't want to mix with your collections management metadata. Namely things that are on a much lower level – where we are talking about the actual files and digital objects that compose the work – for instance, characterization or fixity metadata. I can't imagine a reason to include that with public facing information about a work. It is when we are talking about these more high-level abstract concepts, such as how to display the work, how to properly render it, this is of potential interest and use to the public.

What resources to do you rely on for dealing with the preservation needs of the works- the knowledge that goes into creating that preservation metadata?

Well, I feel that our field is seeing the emergence of a new type of connoisseurship. If you look at digital preservation over the course of the last 20 years, the 1990s were characterized by high-level theory, but not a lot of action. I rather agree with Mathew Kirshenbaum's (UM) that this work requires us to be digital paleographers. Media conservators must be true connoisseurs not simply of the work, not only experts in theory and ethics, but deeply in touch with the material properties of the work. For attaining expert knowledge on materials and technologies present in a given artwork that may be quite obsolete or obscure, I quite honestly rely incredibly on discussion forums and communities buried deep within corners of the web. It would be hard to pinpoint one resource. These days due to the case study on emulation that I am working on at MoMA, I have been spending a lot of time in the MESS project IRC and forums, as well as the Emaculation forums. Fortunately for us conservators, the web just so happens to be the place where people with immense amounts of obscure knowledge of very particular technologies go to share this knowledge with the world.

I think openness is also crucial here. I find that by openly sharing what I am working on, people that may be of help tend to come out of the woodwork. Hardly a novel concept, but perhaps less practiced in major institutions – at least so far.

You mentioned there is the issue of migration vs. emulation and the use of original display platforms, for example, an artist might request that her work be shown on Netscape 3 and you can look up all the technical data related to that platform. However, when you look at the long-term, do you think that will be practically sustainable?

Absolutely. Emulation is far more sustainable than the idea that a museum will simply hire someone to re-code or port a piece of software. This is not a dichotomy - it's not one or the other, but the fact is that emulation is a far more sustainable practice than considering a work's technologies to simply be mutable or variable. Firstly, with emulation, you have one solution that may apply to a variety of works that require the same platform. Secondly, the platforms in question, be they operating systems or the hardware they run on, are ubiquitous to the extent that there are many stakeholders interested in emulating them – for being able to play classic video games, run obsolete scientific experiment simulation software, or to breathe new life into a time based work of contemporary art. Why not tap into this?

The discussion of 'migrating' the emulation strategy, instead of the asset, is fascinating. The change happens on the emulation- and the requirements there- as opposed to migrating the original assets at the artwork level.

7

I think many people don't understand how trivial emulation is to employ as a tool for conservation.

Your use of the word "trivial" is very interesting, could you explain?

It is trivial in the sense that if one is willing to spend the time to learn how to use these tools, there are ample resources, knowledge bases, and communities at one's disposal.

So, the resources are out there for people to find and use.

Yes.

Do you think those resources are going to be there indefinitely?

This is an interesting question. Today it is very easy to find resources on emulating personal computers dating from the late 70s to today. I am inclined to think that these platforms and tools are so engrained in our cultural heritage that this practice of their study will be passed down. For instance, both Basilisk II and Sheepshaver [two emulators], are primarily maintained by two developers that picked up the projects after their creator open sourced the code and ceased development.

Do you think the museums and other cultural heritage organizations that collect these types of works could form a community that jointly maintains some of the more important platforms for the works over time?

Yes, absolutely. I think collaboration and open source software development of commonly needed tools in cultural heritage institutions is an excellent model for sharing the burden.

Are you suggesting that instead of thinking about the preservation of TBMA in terms of individual works, we should think about the preservation of display platforms or environments over time that would apply to whole classes of works?

Yes and no. It really depends on your collection, and the work at hand – this isn't really something that can be generalized. Chances are that in many cases emulation simply doesn't apply to the work at hand. For example, a complex installation with projectors controlled by a computer over a SCSI connection is not something you can emulate at face value. This would require a more hybrid, case-specific treatment. At MoMA however, we do have many software/born-digital works and design objects in the collection that are absolutely perfect cases for emulation, which is why we are currently engaged in a case study that looks carefully at our options here. Mainly what we are considering is fidelity. For each work in the case-study we are assembling what we believe to be an ideal configuration of original hardware, and then studying this in comparison with emulation of

that very specific hardware. When we do things like CRT emulation, we want to be able to point to a specific make/model of CRT and say that we are emulating the characteristics of that specific CRT based on material evidence, not simply first-hand knowledge of what "looks right."

In a different context, Ewan Cochrane, formerly of the National Archives in New Zealand, conducted a similar case-study looking at government documents, and what was lost through the migration process.

That is very interesting to me because so far many of the people that we have spoken have emphasized the works, not the environments.

I don't think that you can separate the two. Even if the platform for a given work is variable, it is crucial to understand the particulars of the underlying platform, so that any mediation or change to the experience of the work is fully understood and anticipated by the conservator.

Are there best practices in the field for emulating artworks?

I would say that there are certainly best practices that I borrow liberally from, that derive from varying areas of practice, but that these really have not been formalized within the art context. I would say this begins with how you are handling the original hardware, the tools you are using to produce disk images, and then eventually, the tools you are using for the emulation.

Could you talk about your current project at MoMA where you are developing a Digital Repository for the artworks? Are you aiming for a Trusted Digital Repository?

The Digital Repository for Museum Collections (DRMC) project was started by Glenn Wharton roughly around 2010. We have an incredible team internally at the museum that consists of leadership from IT, Conservation, Collections and Exhibitions, and recently I've added a few curators and collections specialists. For three years, Glenn worked with this team as well as outside experts to carefully formalize what MoMA's needs were for the stewardship of its digital collections, and what functional requirements would meet those needs. Together with Kara Van Malssen of AVPS, he authored a document that fully articulated these needs, all the way down to use-cases and user stories. When Glenn brought me on to manage the project, I realized that a decent portion of our functional requirements could be accomplished with Archivematica. I was already quite familiar with the software, so we conducted a little pilot to determine if it would suit MoMA's needs. We found it did, and so started working with Artefactual Systems to conduct a formal pilot and some scalability testing in a production setting. Archivematica of course though is nothing

more than a processing pipeline that makes good AIPs. In the end, we really needed something – a tool – that would facilitate proper care, management, and maintenance of the digital objects in MoMA's collections.

We are in the midst of building a management application that will be, the first digital repository system that is tailored to suit a museum – and more importantly – is able to facilitate the careful information modeling and asset management needs of time based media art conservation. This web app will not only facilitate functional collections management tasks such as auditing fixity checks, discovering at-risk artworks, and documenting conservation interventions – but will also allow us to express the variable and iterative lives that these artworks lead. This ability is actually crucial, as various iterations of an artwork (be it an immersive time-based installation, or video game) often possess digital materials that are unique to that particular iteration or expression of the work. As I speak we are working with Artefactual Systems on some of the first user interface wireframes. We are on schedule to have something we can show in early 2014.

Perhaps the most exciting part of the entire project is that we are not reinventing the wheel. We are simply improving on certain existing aspects of Archivematica, and building new function and features into AtoM – the open source access system commonly paired with Archivematica. In the end, this will have the appearance and utility of a completely new and novel tool – but will all be part of the main branches of Archivematica and AtoM, meaning free and open source under AGPL3.

As to whether or not we are aiming for a "Trusted Digital Repository," I would answer that we are aiming for a repository that meets MoMA's functional needs, and properly safeguards the integrity, security, and longevity of MoMA's digital collections. A byproduct of that is that we will check off many of the items on the Trusted Digital Repository Audit Certification checklist (TRAC), but meeting these requirements as a standard is not our goal, and we will certainly not seek TDR certification. While TRAC is certainly useful, it is important to remember that it was created for a very particular context – it was created as a means for certifying compliance with OAIS – a standard which was designed by the Consultative Committee for Space Data Systems. OAIS and TRAC are useful, but it is important to take these standards with a grain of salt, and not prioritize them over the actual on-the-grounds needs.

Many people that I am talking with are interested in the kind of decisions that you are making as you develop your repository. Will you share your decisions as a model for others?

We will absolutely share our decisions and results, and again, the system we are building will be freely available to any institution to use and modify to suit their needs.

To shift gears to training, I wanted to come back to your comment on the shortcomings in the current state of training for working with complex software based works and planning for their preservation. What do you think would be ideal training for the field that you work in?

I think that it needs to be hands-on, and present a hybrid of information science, computer science, art history, and art conservation theory, and ethics. Hands-on, lab environment training with specific cases and works is critical. While there are a few existing masters programs that touch on parts of this, not a single one is sufficient in my opinion. I also question if a program will or should exist that is so interdisciplinary for such a specific and niche field that does not have a surplus of job opportunities. I think one can specialize in one of the above areas and assemble the rest from parts.

I know art conservation programs approach training this way, with hands on work in labs. Do you think this type of technology heavy training could be included in more traditional art conservation programs?

Absolutely – it simply must. Technology is a material used by contemporary artists. Period.

I think that you need a Rhizome apprenticeship program.

Yes, Rhizome actually does have a regular cycle of research fellowships for graduate students interested in furthering their research in digital conservation.

Has has your background as an artist affected your approach to your work as a conservator?

Absolutely. I feel that my background provides me the ability to empathize with the artist on a very meaningful level. There is certain sensitivity and connoisseurship of materials that can only be attained through doing and making. This applies to any area of practice – anyone coming to a work of art carries their own area of expertise. This is precisely why Glenn Wharton established a working relationship between MoMA and Deena Engel, a professor of Computer Science at NYU. About once a year Deena forms a very small group of undergraduate CS students – usually from 2-4, and they engage in the technical documentation of artist's source code. They are able to perform scholarship and research on decisions that the artist made in their code that no art conservator ever could discover without a degree in computer science.

What areas in media art preservation are not getting enough attention right now?

First, web archiving is woefully behind. All of the currently available tools for web archiving are built for the web as it was ten years ago – as a series of pages. The web today is so much more fluid and application based, or stream based. We need tools that account for this.

Secondly, I think that many institutions have not truly come to face the facts that they must act *now* when it comes to the preservation of born-digital or moving image materials in their collections. Anything on video tape or floppy disks over twenty years old should be immediately prioritized. We aren't talking about "someday this will be gone" anymore – we're talking about "this will be gone tomorrow" if you don't do something today.

Finally, just as we must be digitizing our collections for preservation purposes – removing artworks from at risk formats – I think that none of us have truly come to grips with the sheer amount of storage that we need to devote to our collections. MoMA has been great about this, and we allocate very significant increases in storage for the digital collections regularly, but even we have encountered some startling numbers. I think it is feasible that some smaller institutions simply will not be able to afford the storage they need. This is a difficult reality to face.

What do you see as the distinctions among "standards," "guidelines," and "best practices" in the preservation of time-based art? When/if is each most appropriate?

Simply put – guidelines are the base level of what to do, best practices as the vetted ideal way of doing it, and standards as the agreed upon and interoperable way of sharing it.

The first and most important thing is to understand what your institution needs. What are the problems your collections faces, and what can you do within your means to fix them. I can't tell you how many institutions I see blindly trying to adopt or invent some standard, thinking it will solve their problems, yet not being able to demonstrate why they really need it. In practice, carefully informed action that is sensitive to the needs of your institution is more important than meeting a best practice.

Do you think there are standards that are applicable and could be brought in from other fields?

Something I am very interested in at the moment is the application of FRBR to the practice of time based media conservation. This is something we are adopting in MoMA's DRMC. I look forward to sharing how this looks in practice very soon.