Elise D Moore. Towards a Common Schema in Distributed Humanities Research.
A Master's Paper for the M.S. in I.S degree. July, 2010. 91 pages. Advisor: Stephanie
Haas

As humanities scholars move from creating traditional scholarly publications to creating
online collaborative research projects they have begun to realize the need for a common
vocabulary or schema to describe their domains. This paper explores the study of
moviegoing and the process of creating a schema to describe the field of moviegoing.
The purpose is to involve scholars focused on different aspects of moviegoing in schema
development so that they might be able to begin their research with this core schema, and
share their research with each other using this schema as a crosswalk. Ten different
moviegoing scholars were asked to participate in a Delphi study to help define the field
of moviegoing for future research and analysis. The iterative process of a Delphi study
allowed me to collate the thoughts of experts from around the world. It is hoped that the
creation of an after-the-fact schema for existing.

Headings:

      Schema.

      Delphi Method.

      Collaborative Humanities.

      Moviegoing.

TOWARDS A COMMON SCHEMA IN DISTRIBUTED HUMANITIES RESEARCH

by
Elise D Moore

A Master's paper submitted to the faculty
of the School of Information and Library Science
of the University of North Carolina at Chapel Hill
in partial fulfillment of the requirements
for the degree of Master of Science in
Information Science.

Chapel Hill, North Carolina

July 2010

Approved by

_____

Stephanie Haas

Introduction

> "One of the major strengths of the World Wide Web is that virtually everyone who owns a computer may contribute high-value information - the real challenge is to make valuable information be found." (Staab, et al., 2000, p 1)

The World Wide Web is more than a place to search for people and things; it is a place where scholarship and information are published quickly, cheaply, and in new ways. More and more humanities scholars are looking at online technologies and collaborations as potential spaces for scholarship. Humanities scholars have been much slower than their counterparts in science to utilize online collaborations in their scholarship, but they seem able and willing to adapt new tools if they are useful (Baruchson-Arbib and Bronstein, 2007). Work in the humanities has traditionally been a solitary endeavor, however with digital technology providing new ways to present, sort, and view information (spatially, relationally, aggregated, etc.), the humanities field seems to be ready to participate in online, collaborative projects.

The traditional model of humanities scholarship entails long hours in libraries and archives researching with primary source material and producing papers, books, and journal articles, which are, in a sense, a synthesis and extension of the materials that have been uncovered and identified. These papers are then submitted to peer-reviewed journals in hopes that other scholars will find the paper valid and confirm the intellectual authority (Nicholson, 2006). In lieu of writing an academic book or

journal article, many humanities scholars are now creating digital collections. Digital collections are collections of digital materials, curated much like a museum exhibit, with the researcher's arrangement and description building the perspective. By creating digital collections, a scholar has the opportunity to bring together vast amounts of "uncovered" primary resources to illustrate and further expand his/her research and scholarship. With the use of digital surrogates, a scholar is able to put together material that may be housed in disparate locations and make them available to people around the world. By creating digital access to these materials, the scholar has the opportunity to engage other scholars in related fields who might not otherwise have been aware of these resources, as well as other potentially interested audiences.

In order to bring materials online and provide an easy way to search and browse, the digital objects need to be described and a database needs to be created. A database is a structured way to organize information and relationships. This structure is needed in order to perform queries on the data and to facilitate multiple ways to browse and search the content. However, defining a field of study in the humanities through use of a database would formalize the nature of the field in a way that humanities scholars have not necessarily thought about before. (Frenkel, 1991, p 53) Formalizing domains that had traditionally have been studied through abstract concepts by individual scholars pose new sets of questions for scholars. Abstract concepts allow for flexibility in definitions, which has allowed for scholars to craft their various interpretations of ideas. However databases and schemas ask that the scholars not only define specifically for themselves what these terms mean, but in the case of large

group collaboration they need to define these concepts together, with scholars who might have slightly different perspectives on the subject

Many of these digital projects have been, and continue to be, curated and published by museums, libraries, and scholars. These projects are often the product of years of research and many hours of undergraduate and graduate labor to collect, catalog, and present the materials online. Due to different funding sources, support networks, and affiliations, these projects are often developed independently of similar work done by their colleagues. The complex databases behind these projects make integrating data from multiple databases at a later point very difficult. When scholars want to merge data from their projects, differences in how they have defined terms become problematic. Consider the size of a movie theater. A researcher studying early cinema might call a movie theater that is larger than 300 seats large, however in modern times this is a very small theater. Usually researchers look at data through the lens of what they are studying, and often dismiss a more universal view, because it is not important to their research.

As is shown in the rise of massive online collaborations such as Wikipedia and huge group undertakings like the Human Genome Project, a collection of people can create something much larger than any one person. A scholar who works alone to build a digital collection often locks valuable sources inside a closed and usually narrowly focused database specifically designed for a single project and purpose, isolating valuable scholarly content from other possible added value. In order to allow content

from different databases to be linked together without losing meaning, there needs to be some common definitions or ideas by which to connect the data together (Rosenthal, et. al., 2004). There are many different ways one could define or formalize the representation of an idea, depending upon one's particular interest. For example, to someone studying film, a movie theater may be any place in which a film is shown, perhaps because of a focus on the cultural influence the film may have had. However, for the field of moviegoing scholars, a movie theater is very specifically a public place in which films were shown, perhaps because they are more interested in the space as a cultural force, and not the movie itself. By gathering related scholars and projects into "community of interests" (COI), there is a higher chance of being able to develop the generalized definitions needed to make the data interoperable, while maintaining their utility for research (Martin and Groh, 2001). By querying a group of scholars who study specifically the field of moviegoing, I hope to define a core group of terms, definitions and tags that are broad enough to apply to each participant's research, but still specific enough so that there could be a benefit in using this schema as a tool to exchange data across collections.

Background

"Going to the Show" (http://docsouth.unc.edu/gtts/) is a digital library project hosted at Documenting the American South, University of North Carolina, Chapel Hill, which was launched in the summer of 2009. The project was guided by Dr. Robert C. Allen, whose research focuses on the experience of moviegoing in the early 1900s in North Carolina. The collection contains articles about travelling exhibits,

controversial movies, protests, race issues, and all sorts of related information, which could be interesting to people studying other aspects of movies, theaters, city development, moviegoing, urbanity, etc. This collection connects content from North Carolina in new and innovative ways by digitizing and creating a searchable relational database, as well as a mapping interface for exploring the content. The database is focused around the idea of a "venue", being a place where movies were shown. A venue can be any type of building where movies where shown. Venues have specific attributes, and are related to companies, people, and primary sources. The primary sources include records in the Film Daily Yearbook, articles from local newspapers, architectural blueprints, postcards, and photographs. The project was able to link to other already existing content on the web, as well as bring together digital copies of content from all over North Carolina. However, it is still a closed collection, which sits on the servers at the University of North Carolina, and can not be easily combined with other databases or projects about related subjects.

During the course of creating the database for "Going to the Show", other projects related to moviegoing (See Appendix 1: "Going to the Show" Experts Meeting) were identified. The scholars involved in these various projects were invited to a day-long workshop (See Appendix 2: "Going to the Show" Experts Meeting Schedule) to discuss the domain of moviegoing. During the workshop, we covered some of the decisions that were being made to define the "Going to the Show" database, and asked for input from the scholars so that we could begin to define our data in a way that the scholars would find meaningful and perhaps useful to their own scholarship.

Each of these scholars approached the domain of moviegoing within their digital projects from slightly different angles, often due to the specific interest of their lead scholar, but other times due to technological constraints or choice of tools. For example, Jeff Klenotic, a professor from the University of New Hampshire, described his "moviegoing database" to the group, which was run entirely on his computer in an instance of ArcGIS. He had taken a GIS class at NASA, and began using the tool for the research of moviegoing in New Hampshire. ArcGIS is a very extensive GIS package, which allows him to do complicated spatial analysis, which many of the other projects did not focus on.

The scholars invited to the discussion were very interested in finding a way to share their data with each other. The overlap they recognized suggested that it would be useful to explore ways of importing and exporting data between projects in order to allow comparison and other types of analysis. For example, Jeff Klenotic, who was also planning research in New Hampshire similar to that of "Going to the Show", was very interested in the comparison of movie theaters in similar sized cities. Being able to combine his research in New Hampshire with that of "Going to the Show" in North Carolina might allow him to reach broader conclusions than would be possible with only his own research. In order to merge data together from various sources, concepts and relationships need to be mapped to similar concepts and relationships. The database schema represents these relationships and the data elements within the schema, or the terms represent the concepts. A schema in psychology is a mental structure that represents some aspect of the world. In database terms, a schema is the

structure of a database system described in a formalized language. The data elements within the schema are defined in both technical terms and conceptual terms.

Research Questions
The overarching research question I seek to answer is "What challenges arise when a community of interest tries to create a common schema to facilitate the aggregation of information?" Specifically, I am examining concepts and vocabulary needs of the domain of moviegoing as well as the schema, or relationships among these concepts.

Importance of Study
While the ideal approach would be to create a common schema and terminology for a Community of Interest from the beginning (Rosenthal, et. al. 2004), this is not always possible. Scholars in the humanities are moving from individual research projects to larger online collaborative projects and the possibility of combining their research could produce interesting results and juxtapositions.

When creating a database, one must define the data model, or schema that is to be used. This data model is the organization of various concepts and their relationships to each other. If all of these scholars could work within a shared schema, comparison of data would be easier. Using the same database, or at least parallel development of projects could have allowed the scholars to develop a shared data model from the start. However, as mentioned before, varying funding, timing of projects, proximity of the scholars and other resources has guided their development of separate databases.

In this study, I asked the members of this scholarly community themselves to help define a base level of terminology or definitions within the field of moviegoing, in order to be able to effectively share data across collections in a way that would still be useful for their research.

What is the landscape?
A "Community of Interest" is a community of people with a shared interest or passion, and is one way to provide a common ground even if other organizational aspects are lacking (Rosenthal, et al. 2004). The community of moviegoing scholars is a good example of a group with a shared interest but with no overarching organizational structure in place.

In Rosenthal, Seligman, and Renner's paper "From Semantic Integration to Semantics Management: Case Studies and a Way Forward", the authors explore the shift from after-the-fact semantic integration to managing the semantics of a system from the start. After a few years of trying to create a single data model for their entire company, they realized that this task was too herculean, as there were too many different focal points to be able to create a realistically usable model. They state that although a single data model might seem most ideal, they have learned through previous experiences that semantics should be managed "within communities of interest (COI) rather than standardizing data elements across the entire department" (Rosenthal et al 2004, p. 47).

> According to Rosenthal, et al. (2004):
> "… we must go beyond after- the-fact semantic integration among existing systems, to actively guiding semantic choices in new ontologies and systems – e.g., what concepts should be used as descriptive vocabularies for existing data, or as definitions for newly built systems." (p. 44)

Despite the differences in how researchers have begun to describe their domain through databases and schemas, the specialized nature of this domain (moviegoing), seems to be fertile for the development of a shared ontology and detailed data concepts. Bringing this type of data together from different projects with varied focus, and interests, relies upon a common understand of core terms, definitions, and concepts which are of interest cross-collections.

Lessons to learn from the Sciences

The history of the Human Genome Project[1] provides an example of how independent researchers with shared interests worked together as a community of interest to accomplish their shared goals. In the early 1990s, according to computational biologist Tom Marr, of the Cold Spring Harbor Laboratory in New York:

> "Biology [was] in a phase in which it's going from small, single-laboratory investigators into the computer age because instruments are pushing it." (Frenkel, 1991, p 45)

One of the first big questions with the Human Genome Project, much like the questions driving this project, was how to create a shared data model to encompass data from pre-existing projects. Before the project was envisioned labs defined relationships between elements as suited their needs for the individual projects they

---

[1] http://www.ornl.gov/sci/techresources/Human_Genome/home.shtml

were working on. It was soon discovered that it was not a common data model between projects which matters when moving to merge databases, but the concepts which are used need to be common across all of the databases (Frenkel 1991, p 50).

Genome sequencing began in small labs, which were attempting to understand one particular part of the genome in order to develop new medicines or further understanding into a particular medical problem. With private financing this was the most logical approach, however "…the cost of finding individual disease genes, one at a time, is often staggering" (Cantor, 1990, p 51). In the mid 1980s the Department of Energy (DOE) and the National Institute of Health (NIH) separately began pursuing the goal of mapping the entire human genome. After a "14-month-long deliberation [that] led to a unanimous report, which urged the United States to begin the Human Genome Project and to work cooperatively with other nations who wished to jointly pursue the common goal" the DOE and the NIH combined their resources and began the Human Genome Project (Watson, 1990, p 45). The Human Genome Project (HGP) contained geneticists from the United States, United Kingdom, France, Germany, Japan, China, and India. The HGP began by laying out a plan to move forward with the project, focusing on how to merge already existing data, in order not to re-map any part of the genome that had already been identified. Beyond combining existing data, the HGP planned the mapping of the remaining parts of the genome.

It was "unlikely that these [sections of the genome] can be effectively parceled out as assignments, it [was much more] likely that individual interest groups would evolve

into effective working partnerships for each genome region" (Cantor, 1990, p 51).

The question of how to store this information was in the forefront of everyone's

mind. It was understood that a single genome database would not suffice for the

distributed nature of this field, besides each lab had its own unique needs for

presentation and analysis of the data. Also, each lab had varying degrees of resources:

"A lab might go with relational technology simply because the more mature

technology could make its staff less dependent on programmers" (Frenkel, 1991, p

50). This again posed the problem of integration.

> If there are to be many separate genomic databases, some relational and some
> object-oriented, will data from each be comparable? (Frenkel, 1991, p 50)

Frenkel states that the problem does not lie in the database type or structure, but in

"the concepts that are used in creating the database" (Frenkel, 1991, p 50). The

Human Genome Project accomplished the enforcement of shared data concepts, by

developing data sharing guidelines, which made a condition of the genome grants the

sharing of data after a six month period- in order to allow for checking of data and

development of drugs or other inventions (Human Genome Program, 1993). This

stipulation encouraged the use of similar data concepts by each institution in order to

facilitate the sharing of data in various gene banks, that had preset data models.

The process that the Human Genome Project went through was very long, but in the

end they were able to map almost the entire human genome five years ahead of

schedule. The Human Genome Project is a model for distributed work and

combination of resources. One of the main struggles for the HGP was the

formalization of authority, concepts, and mapping of ideas in a way that was broad

and flexible enough to change, and still be useful as the research evolved. (Frenkel, 1991) This model of flexibly defining semantics of a domain in order to allow the rapid development and release of information has been a model for many fields, from other scientific endeavors (Brooksbank and Quackenbush, 2006) to ones within the humanities (Ball, 2010).

There is much that humanities scholars can learn from their science counterparts. Despite the radically different topics, the group of moviegoing scholars queried for this project have a lot in common with The Human Genome Project. They wish to find a way to share data within their domain, without stifling the workflow and projects already in place occurring around the world. The excitement and concern are also shared. Hopefully by following in the example of the Human Genome Project by working directly with scholars to define their field, this project will be able to create a common semantics, which will allow the future merging and comparison of data.

Method

In order to come up with shared definitions, I molded an experiment using the Delphi Method, in hopes of encouraging participation by the very specialized scholars of moviegoing from around the world. The Delphi Method allowed me to systematically elicit ideas from scholars in disparate locations over email, and help the group come to a consensus.

Method Justification

 "The Delphi Method recognizes human judgment as legitimate and useful inputs in generating forecasts. Single experts sometimes suffer biases; group meetings suffer from "follow the leader" tendencies and reluctance to abandon previously stated opinions" (H. Murat Gunaydin, 1998, para 2). The Delphi Method's systematic approach and ability to build an anonymous consensus of experts around the world makes it an effective tool to solve complex problems such as defining a schema or collating strong opinions. In order to create consensus on the design of a common schema about moviegoing and definitions of relevant terms, a group discussion needs to occur. However, due to the specialized nature of the sample field and the diverse geographic locations, a group meeting would be too difficult (and expensive) to coordinate at this time. The Delphi method allows each participant to think about the concepts on their own time and be a part of this iterative consensus building process.

Overview of the Delphi Method

"Project Delphi" was the name of an Air Force-sponsored study, starting in the early 1950's, that tried to "obtain the most reliable consensus of opinion of a group of experts" (Dalkney and Helmer, 1963 p.458) about possible Soviet attacks on United States soil. This new approach was used because of the expense and time of other methods (Linstone and Turoff, 1975 p.10). In 1963, Linstone and Turoff wrote a seminal book about this process, *The Delphi Method: Techniques and Applications*. This book outlines the basic parts to the Delphi Method, and shows examples of studies that have used the approach. Linstone and Turoff's book is the major resource for my commentary about the Delphi Approach.

There are many different interpretations of the Delphi method, but it breaks down to four basic parts, which can be iterated until a consensus is reached.

1. The first round of the Delphi Method begins with an open-ended questionnaire. This allows the moderator to get a feel for each of the participant's responses to the information space, and is used in order to gather perspectives for review in the second round.

2. Each participant is given a second questionnaire that contains responses from the first round summarized by the moderator. Each participant is asked to again discuss any discrepancies. This round is the beginning of consensus building, and also where disagreements will begin to become evident; as it is the first time the participants are able to comment on each other's opinions.

3. In the third round the participants are again given the consolidated and summarized version of the previous round. This round allows participants to make further clarifications and refine their judgments.

4. In the final round, participants are given another opportunity to refine their opinions, but mainly this round gives them an opportunity to see if they have any major objections to the final group consensus.

There are three key characteristics of the Delphi method that distinguish it from other methodologies: structured information flow, feedback and refined opinions from the participants, and anonymity for the participants. The feedback from the participants allows them to adjust their definitions, in order for the process of reaching a consensus to take place. Anonymity is a very important part of the Delphi processes.

Because "anonymity [is] employed, each participant [is] forced to judge the potential of each possibility on the basis of his [or her] knowledge and the supporting arguments presented" (Linstone and Turoff, 1975 p.189). This helps to curb the "band wagon" effect, which is often seen during group discussions.

As noted previously, there are many applications of and benefits to the Delphi method, particularly providing a systematic way to explore a subject where other qualitative methods do not apply. The limitations of the Delphi method can be minimized if careful thought is put into preparing the study. Selection of participants is one of the most important steps in the process, because it directly affects the results (Hsu and Sandford 2007, p. 3). The quality of the subjects is emphasized, however the number of subjects needed to reach a consensus is not agreed upon. Time requirements are another major obstacle to the Delphi method. Since the surveying is done from a distance, each of the rounds can take up to two weeks, or more.

If performed with a group of experts and with relatively low bias from the moderator, the Delphi method has been proven to be fairly reliable. I have the potential of bias on the subject of moviegoing because of past work with "Going to the Show". However I hope this knowledge helps in understanding the various projects and scholar's needs, and does not prove a hindrance in defining specific terms and data structures.

Participants
Ten university professors from around the world in the fields of Communication Studies, American Studies, Cinema Studies, and Theater Studies were invited to

participate in this study. Eight of the ten agreed to participate in the study (See

Appendix 3: Participants of the Delphi Study). They were identified based on their

interest, and often-current development of digital projects relating to moviegoing and

cinema culture. This is not an exhaustive list, but each one is an expert in the field,

which is one of the most important stipulations of the Delphi process. Subjects

participated voluntarily without compensation. The eight experts who elected to

participate in this Delphi process were from America, Europe, and Australia. Their

common language is English, however it is not a first language for all participants.

There are also differences in the culture of the participants, which may affect the

research they focus on. For example, in the United States, there is much focus on

racial segregation, and the eventual racial integration of movie theaters. In contrast, in

Holland there is much less focus on race, but more interest in political association as

movie theaters were often used as vehicles to promote ideologies. These differences

may help this study to define the field of moviegoing in a general way, so that one

perspective is not favored over the other.

Initial Materials
The research project began with the collation of two moviegoing databases, the

Cinema Context database (See Appendix 4: Cinema Context Entities and Appendix 5:

Cinema Context Schema) and the "Going to the Show" database (See Appendix 7:

"Going to the Show" Abstracted Schema). I started with these two databases mainly

because of access, but also because these are two established digital moviegoing

projects, which have already worked through a lot of the initial questions in defining

this field. I merged the two database diagrams relying upon my knowledge or

interpretation of the fields listed in the diagrams. Not having the field definitions in front of me, or access to a data dictionary, for the Cinema Context database perhaps led me to misunderstand some of the fields. This is part of the reason the definitions for the fields were initially left blank, so that the participating scholars could help tease out any ambiguity.

 "**Cinema Context** contains records of all cinemas that have existed in the Netherlands since 1900. You can find information about the management, the building's history or the architect. And if a travelling cinema passed through your hometown a hundred years ago, Cinema Context usually can tell you more about it" (Dibbets, n.d.). "Cinema Context" goes into detail about specific movies shown at theaters, as well as traveling exhibitors, whereas "Going to the Show" focuses more on place, and what was associated with that space. Having access to these two database schemas allowed me to create a data model that was slightly broader in coverage than one based on a single project.

The "Going to the Show" database also included tags defined by the project's scholarly advisor Dr. Robert Allen. These tags were used to help describe articles and advertisements found in various newspapers about movie venues in the early 1900s in North Carolina, as well as the venues themselves. They were created as another way to group, describe and provide access to the bulk of primary source material scanned for the "Going to the Show" project. The tags from "Going to the Show" were used as a starting point for the tags used during my Delphi study.

Process

 The results of the Delphi Process were analyzed after each iteration. As the

moderator, I tried to be careful not to impose preconceptions or over-simplify the

concepts being explored. While the participants were provided with an initial set of

terms, their definitions and uses may not always be the same. For example, what is

the place where moviegoing occurs? Is it a cinema, a theater, or movie venue? A

movie venue could mean any place where a movie is shown, where as a movie theater

is usually a building specifically created for the purpose of showing movies, and a

cinema, is similar but comes from England and is a term that is not used in much of

the world. These variations in vocabulary make a difference in the type of

information that gets collected. According Dr. Robert Allen, the term cinema "could

NOT be used to handle movie exhibition prior to around 1905 because there weren't

any cinemas (despite the fact that millions of people saw movies)." While this study

did not search for one definitive answer, the study would be successful if the resulting

schema and field definitions could be used by a majority of the participants in their

individual research. Another sign of success would be if participants and others who

study moviegoing in the future continued the discussion initiated by this study.

It is understood that not all scholars were able or desired to share their research with

others. It was also important to try and maintain anonymity through out the project.

Participants may become aware of the other scholars involved, but it was the

responsibility of me to ensure that the responses to various Delphi surveys remained

anonymous during the course of the study.

In round one of the Delphi process, participants were asked to review the combined database described above, and consider if the combined schema could satisfy their research needs. Participants each received an entity relationship diagram (See Appendix 9: Initial Combined Schema (Delphi Round 1)) of the combined database, a description of the database fields in a simplified data dictionary (See Appendix 10: Initial Data Dictionary (Delphi Round 1)), and a link to the online survey of tags (See Appendix 11: Initial Tag Survey (Delphi Round 1)). Participants were asked to mark which of the fields they felt were essential fields for their research and should be in the shared schema. Further, they were asked to provide examples or explanations of how they have used this type of information in their research. Participants were also asked to select a set of relevant keywords from the set of tags from the "Going to the Show" project. This round was intended to help develop a common vocabulary to further describe the concepts in the field. Participants were given the longest amount of time for this first round of analysis, as the initial introduction to the project was expected to take the most time.

Participants' responses were analyzed and merged with the original database and list of tags, preserving any field or tag which was noted as important by more than one person, and adding any additional field or term mentioned (See Appendix 12: Combined Schema (Delphi Round 2) and Appendix 13: Data Dictionary (Delphi Round 2)). The survey was modified to show definitions and examples that were given for agreed upon fields, and notes about differences of opinions on disputed fields (See Appendix 14: Tag Survey (Delphi Round 2)). The comments were edited to help maintain the anonymity of the participants' answers.

In the second Delphi round participants were asked to review the new fields and mark which ones they felt were essential fields to include, and to look over any of the fields where there was disagreement, and try to comment or further explain their opinion. This round is an important step in consensus building as it is the first time the participants see each other's input and begin to refine the collective thought about crucial and non-crucial elements as well as add any additional possibly useful tags to the list.

I again collated the responses, and sent the information back to the participants for a third and final review of the terminology and list of tags. In this round the participants were asked to denote changes they felt strongly about. They reviewed the information one last time. This was also an opportunity for the scholars to provide additional feedback on anything else they felt was missing from this schema.

Results

Ten participants were invited to participate in the study. Nine agreed to participate, but only eight participated in the first round. The second round only had seven participants. During both rounds not all participants responded to every question; many were skipped over.

The first and easiest set of terms Idiscuss are the terms that were agreed upon in round one. These were terms which were universally seen as important from the very

beginning. The next sets of terms I discuss were terms that there was some dispute about in round one, but which was resolved during round two. Some of these terms were added during round one and then immediately agreed upon, whereas others needed a second round in order to bring more clarity to the definition. The third set of terms discussed are terms, where there has yet to be a clear resolution on their definition or relevance. This was sometimes due to the fact that I only held two Delphi rounds, therefore issues brought up during the second round never had time to fully develop, but often because they are very nuanced issues that need more discussion from perhaps a larger group, in order to really understand.

Table 1: Results of Delphi Round 1 Data Terms					
	Yes	No		Yes	No
Cinema/ Traveling Show			**Program**		
Address	8		Date	7	
Opening	8		Title	6	1
Closing	7		Alt Title	5	2
Seating	7	1	Distribution Company	5	2
Screens	5	3	Category	4	3
Performance Type	3	3	Distribution Employee	4	3
Racial Policy	3	5	**Film**		
URL	1	7	Title	8	
Company			Date	7	
Name	7		Country	7	
Start Date	7		Category	5	
End Date	7		Directed by	5	2
Info	7		Length	4	2
Job	7		**Film Episode**		
Alternate Name	4	2	Title	2	3
Employee			Year	1	4
Name	7	1	Length	1	4
Biography	6	2			
Job	6	2			
Birthday	3	5			
Place of Birth	3	5			
Death Date	3	5			
Race	3	5			
* There were 8 participants for this round however, not every participant responded to each field					

Delphi Round 1: Results

The first Delphi round helped to shape the discussion which continued through out

this process by highlighting which elements were accepted by all, and which elements

needed more attention. After the first round of the Delphi process, I eliminated fields

that were marked as important by fewer than two people, as this indicated that the

field was too specific or otherwise deemed unnecessary for a general schema.

Cinema & Company
Most of the participants immediately agreed upon the core elements of the Cinema

table, Cinema: Address, Opening, Closing, Seating, etc. This is the place where

moviegoing occurs, and therefore one of the central elements in almost all of their

research. The majority of the scholars also saw a need for most of the attributes of

Company. This may signal the importance that companies have when analyzing the

history of moviegoing development.

Film Episode
One whole table of the database was eliminated during the first round, the film

episode. A film episode refers specifically to a phenomenon that occurred in early

cinema. A "film episode" or "film serial" refers to short clips usually shown before a

feature film, which were broken into chapters, or episodes. Each episode would

generally be screened for about a week. While this phenomenon continued through

the 1950s, they were most prevalent in the silent film era. Although some of the

scholars worked in early cinema, with this group of scholars there was not enough

universal use for this information to keep it as part of the core database fields.

Lesser Entities: Program, Film

Within most of the other entities (Employee, Program, and Film), the participants had needs for different levels of detail. For example, only about half of the participants thought detailed information about employees such as date of birth, place of birth, and date of death would be important to their research. Philippe Meers felt that this information was useful just for "situating a person", however, Daniel Biltereyst notes, "We didn't work on this for the large longitudinal project (Flanders, 1920-2008), but we are gathering data of this type for the Ghent Cinema Capitol project. But family histories (as in the Capitol case with one family controlling this major cinema network in Ghent) and network histories are often very complex." Most other participants felt that this might be useful information when found, however this type of detailed information is not easy to come by in all cases.

Within the Program and Film entities the needs of the researchers varied as well. For instance, many of the scholars felt that details about a film or program were better left to other databases that focus specifically on films such as the Internet Movie Database (IMDB). Others, however, need film genre and length information in order to determine the shift from shorter serials to full-length feature films. Names and dates remained core elements of every entity.

Participants' quotations were used to build definitions for the terms for which there was a consensus on importance in the first round. By pulling direct quotations from the participants' responses to useful fields to build definitions for each term, I sought to avoid filtering their perspectives through my interpretation.

Delphi Round 2: Consensus

Cinema: Performance Type
Some of the more interesting points came out in cases where the participants did not

completely understand the term; one such example was "Cinema: Performance

Type". In "Going to the Show" this field was used to indicate the different types of

performance, which might occur within a cinema, such as illustrated song, movies, or

dance. However this again is a phenomenon that occurred mainly in early cinema.

The following two quotes from the first Delphi round show the differences in how

this field was perceived and opinions on whether it should be an attribute of a cinema

at all.

> "To date I have used this information primarily as it pertains to the presentation of various forms of sound pictures during the silent film era and also as it pertains to the presentation of local amateur contests."
> (Jeff Klenotic)

> "It's not clear to me what the meaning of 'performance type' is. It sounds like a label to be applied to a building, a venue. Our database mainly catalogs movie theatres as buildings with a special purpose or function. So we probably have collected only one performance type: cinema. … Other venues like town halls, music halls, circuses, bars, etc. are only mentioned as an address (location, property, lot) in relation to a programme. We think that a function or type will become clear by looking at what happens in the venue. Wherever a movie is shown, it counts as a venue, and becomes part of cinema history, even when it is not seen as a cinema proper. 'Performance type' is always temporary and it needs dates in case of type changes. In my opinion, it might be sufficient to mention a performance type in a general note describing the venue."
> (Karel Dibbets)

During the first Delphi round, others noted the importance of tracking different types of performances in order to understand the movement from a mixed performance venue to more focused cinemas that only show films.

> "Key data for charting the movement from music hall or fairground to fixed-site cinema."
>
> (John Caughie)

> "I have only touched on this in my research, but I think it's crucial to understand the cinema as an (often) "mixed" venue, and to be able to track how that changes over time."
>
> (Arthur Knight)

Based on these comments, and noting that the performances that appear within a cinema can be determined by looking at what types of programs (another existing entity) are associated with the cinema, the second round question about Cinema: Performance Type was *"Is this field necessary if "programs" denote whether they are "mixed programs" or not?"* By building this information into another entity, "program", that has a single date of occurrence, the type of performances at a cinema becomes a temporal element of data, and it allows the value to change over time for a specific venue. However, in order for this to be useful,the semantics of the basic database diagram need to be agreed upon. A film or performance is arranged within a program, which plays at a cinema.

Address Specificity
During the first Delphi round it became clear that Cinema: Location was very important for each of the scholars, especially since many of them were using geospatial tools to help understand "geographical, social, class and other parameters" (Daniel Biltereyst) of a cinema. For the second Delphi round, I took the opportunity

to ask the scholars what level of detail of geographic information is important for their research. There was consensus that while GIS coordinates are ideal, gathering them is very time consuming. In their absence, having a street address would be satisfactory. However it was noted that any type of geographic data is helpful if ascertaining a street address is not possible.

Sources

One interesting topic that was brought up by two of the participants was the exclusion of sources from the database.

> "Our database has a facility to add information about sources and publications about a venue or other major entities. We don't collect sources or publications, but we can provide some references."
>
> (Karel Dibbets)

> "There must be a field to capture the reference to the sources where the cinema is mentioned."
>
> (Maria A. Velez-Serna)

For the second Delphi round I responded by stating:

> "Source material for films, cinemas, and companies will hopefully exist for many of your databases. I have opted to leave this out of the schema. I do think that they are important, and should be included whenever possible, but when merging databases, I think it's best to deal with just the "data", and allow the sources, to be browsed, or exchanged through a separate process"

The responses were:

> "I would like to point to a difference in observing e.g. oral history accounts. In your interpretation, these are "sources" but we consider them as "data" of another kind: they supply information on image, distinction, meaning making, etc, that are not available from the hard "data"."
>
> (Philippe Meers)

> "Maybe – but I hope that does not encourage us to 'lose' the reference info. We need to make sure that it is stored consistently and can be retrieved on request."
>
> (Maria A. Velez-Serna)

Participants were concerned that reference information would get lost if it were not specifically included as part of this database outline. At this point, I did choose to interject into the discussion. While the moderator of a Delphi study should allow an open discussion in order to elicit opinions, it is also important to keep the discussion within scope. I felt strongly that this was more of a technical discussion, and not part of conceptualizing the schema. There are accepted standards[2] to record information about references, such as oral histories, newspaper clippings, and photographs. I choose not to focus on this, because there are already many different tools to help allow these standards to be used and to be interoperable with each other.

Open Issues

Cinema: Seating Capacity
During the first Delphi round participants were asked the importance of the seating capacity of a cinema, and if it were best measured in "numbers, or S, M, L?" Almost all of the participants noted the importance of seating capacity to their research and commented that they preferred precise number of seats – however the availability of this information depends on the quality of the data (Daniel Biltereyst). During Delphi round two the participants were asked if a scale would be useful (0-150, 150-300,

---

2. There are many different standards for citing resources, and specifically primary resources – such as APA, MLA, Chicago Style. The Library of Congress has a good resource on using MLA and Chicago to cite primary resources (http://www.loc.gov/teachers/usingprimarysources/citing.html).

>300). In response to this, all of the participants agreed that scale would work better for the type of data that existed, however the ranges of the scale were not agreed upon. This will be important to explore with a group of moviegoing scholars, as the scales at which cinemas operated varied greatly between different time periods and locations. For example, a "large" cinema in the early 1900s might be one with a seating capacity of 300, whereas in modern terms this would be considered small, or medium at most. With further discussion, a seating capacity scale that is meaningful to all participants can be developed.

Cinema: Racial Policy

Cinema: Racial Policy was the focus of interesting discussion throughout the Delphi Process. This was a field that again, was pulled from "Going to the Show", a project that focused on early cinema in North Carolina where race was an issue in relationship to policy and ownership. There were a few other scholars, mainly in the United States, who used race as an important component of their research. Arthur Knight's research is about African American moviegoing, so this is a crucial field for him. Most of the other scholars noted that while this might be an important bit of information for other researchers, this was not relevant in their location or time of focus.

However there was another type of association that was not covered in the original schema; Political and Ideological association.

> "Racial policies like in the USA did not exist in Dutch cinemas and theatres, although it certainly was part of cinema culture in the Dutch colonies of the East and West Indies before 1940. A religious and social

compartmentalization or Apartheid has characterized Dutch society up to the 1960s. One can differentiate between catholic, protestant, socialist, and 'neutral' attitudes towards the cinema, affecting cinema programming. This is less a feature of a building, and more of the cinema owner and his programming. The role of Jewish cinema owners is the topic of a forthcoming PhD study."

(Karel Dibbets)

Philipe Meers wanted cinema descriptions to include "Political orientation or not; Precise political orientation". Similarly, Daniel Biltereyst would like to link cinemas to "ideological groups and other companies." These comments and interest brought me to ask about the usefulness of Ideological/ Political organizations in the second Delphi round.

"Although this information has not been essential to me thus far, I would choose to include it in the database. For instance, I have often sought to research audiences in order to determine what social class was most likely to associate at/to a given cinema, and although I see this as a broad sociological designation rather than any sort of "official" ideological association, I could see specific class based organizations (e.g. Labor Unions) that ran cinemas as a kind of "ideological association.""

(Jeff Klenotic)

"Ideological association is part of our research, but it is not included in the public side of our database."

(Karel Dibbets)

"I wonder if this category could be broadened, so it could include, say religion (films have often been shown in churches) or social club (Masons, etc.)-that could be useful."

(Arthur Knight)

Arthur Knight's comment about broadening the field to include religion begged the question of what the purpose of this field would be. Would it be used to represent potential for discrimination and favoritism in a cinema for a certain type or kind of film, in order to push some larger social agenda? If so, would racial policy fall under

this same category? Or are they completely separate concepts that need to be handled

by unique fields? Unpacking these concepts and how they relate to the field of

moviegoing may reveal areas of similarity as well as difference. This may help decide

how these will be represented in the database.

Additional Relationships

During Delphi round one, Karel Dibbets wrote out the semantics for links between a

company and other entities within the database.

> "A job is expressed by a link between the company and a venue, a movie,
> another company, a person, an address, etc.. Some information has to be
> added as to the meaning of the link. My suggestions:
> A link with a venue means exhibition by a cinema chain.
> A link with a programme means exhibition by a travelling cinema.
> A link with a movie means distribution (production being outside this
> domain).
> A link with another company means a mother/daughter company.
> A link with a person means employment or ownership.
> A link with an address means ownership of the building located at the
> address."

The first three statements were already represented in the schema. During the second

Delphi round, the participants were presented with these relationships and asked:

> "The following was suggested as a way to indicate the relationship of a
> company to various other entities within the database. The first three
> connections exist within the current database structure. The last 3 are not –
> please discuss the importance of the different relationships, or if you would
> prefer to indicate them in a simple text field related to each field."

There were varying responses within Delphi round two to these new relationships

Table 2: Responses to Various Proposed Database Relationships			
	Company - address = ownership of the building located at the address.	**Company - movie = distribution (production being outside this domain).**	**Company - another company = mother/daughter company."**
Philippe Meers	this need not be the		this is not exhaustive,

	case: a company can rent a building?		as it also could be a sister company, under the same mother
Maria A. Velez-Serna	You mean a link of a person with an address? As the coordinator indicated above, this would be covered by the Person – Job – Cinema relationship, where 'Job' is owner, lessee, etc.	This is related to the 'Programme' issue. What is distributed – a film or a programme? There might be different answers to that. So if possible I would like to see a Company – Distribution – Film relationship.	This would be very useful, especially if these relationships could be built across databases, with company dependencies established around the world.

These statements show that there needs to be more discussion aimed at the specific semantics of the database. Dibbets brings up some possibly important issues about the relationships between a cinema or company to an address, the varying relationships that a company could have with another company, and the different relationships a company or person could have with a movie.

Geography as Place
Jeff Klenotic made the point that the geography of a cinema location could be defined as more than just a set of coordinates or street address, but also as a social concept.

> "First, with regard to geography, the main category of spatial information in this table on "cinemas/travelling shows" is address. Address is a crucial starting point, but I might wish to embed such addresses in broader spatial contexts. "Neighborhood" is one such context. Cinemas are places, but oral histories often reveal that cinemagoers identified these places as part of a larger patch of geography such as "downtown" or the "central business district" or the "west side" or the "immigrant quarter," etc. An oral history participant may say "I remember going to the movies downtown" and they may identify one or more theaters as being part of that area and they may think of those theaters as somehow different and distinct from theaters in other neighborhoods or parts of town. To be able to search the database by such a category as "neighborhood" (or some similarly geographic category) would allow us to pull out those theaters and explore their relation to different companies and to different kinds of moviegoers…"

This is a very interesting topic for many of the moviegoing scholars who are now looking at moviegoing through spatial and social lenses. Unfortunately, because this was brought up during the second Delphi round there was not time to pursue a discussion about the use of geographic names in this Delphi process.

Diagram 1: Final Database Schema
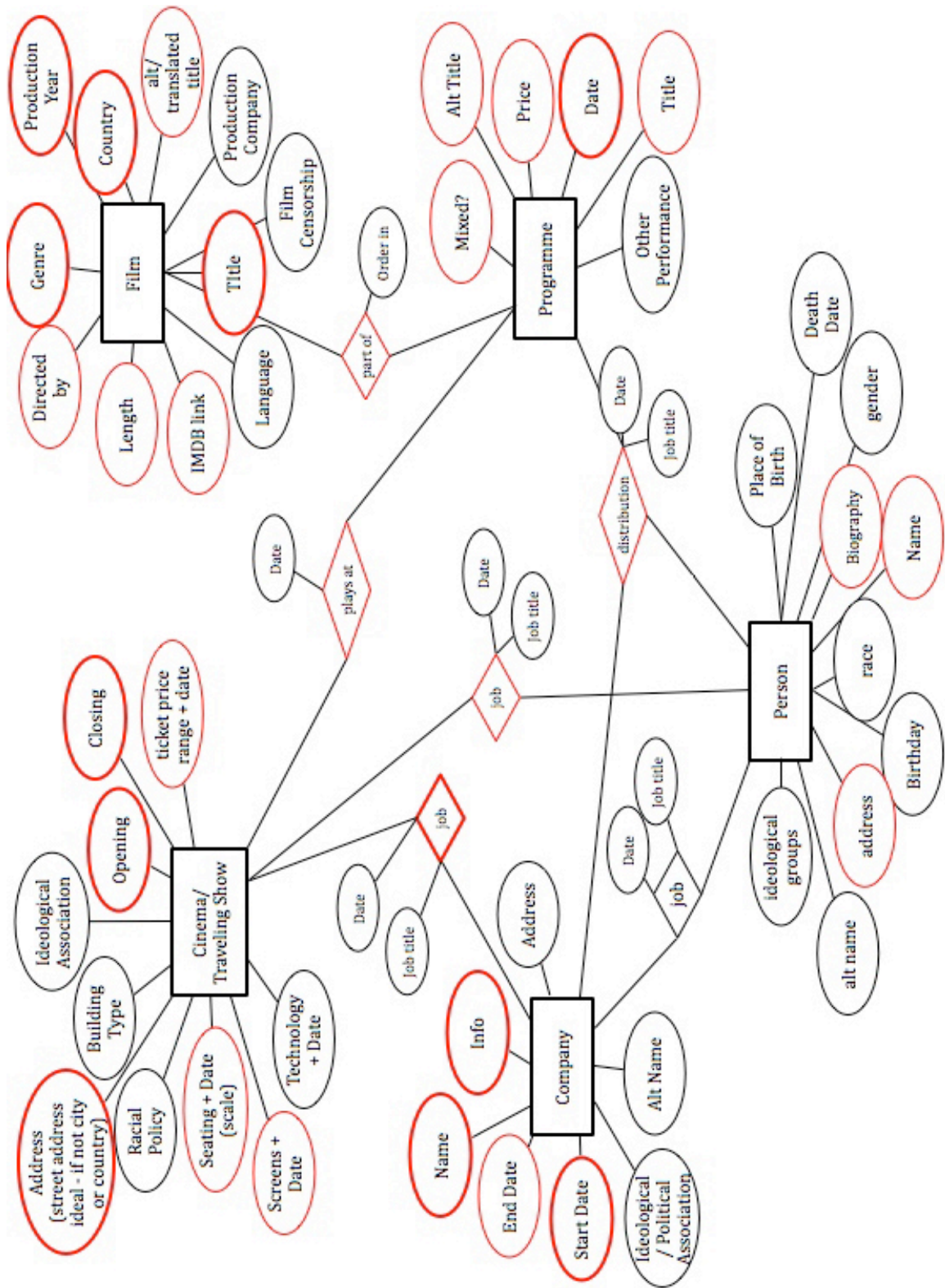
Fields circled in red are core, bold red fields are core fields to almost all projects.

Tags

Tags were pulled from the "Going to the Show" database and scholars were asked to

mark any tags that they think might be of use to them, as well as note any additional

tags which they would find useful. The purpose of this portion of the study was to

find a common set of terms by which projects could begin to tag cinemas, content, or

other items within their database. Tags are useful in digital collections as they

facilitate more access points for users to search and browse the material by. If

scholars used a common set of tags, when they attempt to merge data, the tags would

still be meaningful.

The first Delphi round had eight participants, but only seven participants participated

in the online survey of tags for Round 1, and Round 2 only had 6 participants.

Appendix 11: Initial Tag Survey (Delphi Round 1) shows the list of all the tags

presented to participants during round one. "Votes Round 1" are the number of votes

which each tag received during the first round. The third column shows the number of

votes each tag received during the second round. There were 93 tags in round 1 of the

tag survey, seven of them received fewer than two votes, and were omitted from

round 2.

Table 3: Results from Delphi Tag Survey					
Tag	Votes Round 1	Votes Round 2	Tag	Votes Round 1	Votes Round 2
change in owner	7/7	6	special showing		4
children		4	tent	5/7	3
closing		6	vaudeville		3
company		6	actualities	4/7	5
distribution		5	balcony		4
film title		6	balcony, racially segregated		2

Tag	Fraction	Votes
hours of operation		4
licensing		4
local films		5
manager		4
music		5
name change		6
non-theatrical venue		5
pianist		5
program change		6
purpose-built movie theater		4
renovation		5
seating capacity		6
admission price		6
architect		5
band or orchestra		5
censorship		4
church		5
fair		3
family		4
feature film		6
illustrated song	6/7	2
lecturer		3
location change		4
organ		1
re-opening		6
religious films		4
religious objection		5
storefront theaters		3
traveling exhibitor		3
advertisement		5
charity		4
circus		2
fire		2
ladies and children		3
men only screenings	5/7	1
open-air screening		4
projection problems		1
promotions		4
proposed theater		4
seasonal exhibition		3
sound films		3

Tag	Fraction	Votes
carnival		1
concessions		3
midnight shows		3
newsreels		5
opening notice		2
public appearance		2
school		3
separate entrance		2
social club		4
women only screening		2
WWI		3
air conditioning		3
amusement park		2
black theater		2
color films		4
contest		3
drugstore		2
fashion show		1
minstrel shows		2
morality	3/7	3
opera house		1
penny arcade		1
projectionist		3
racial policy		2
serials		1
streetcars		2
ventilation		2
ymca		3
airdome		2
american indians		1
hotel	2/7	2
motion picture party		2
notable		0
crime		
fight films	1/7	
influenza		
alias		
birth of a nation	0/7	
boxing films		
photographer		

There were 45 new tags in round 2 that were added by the participants during round 1. Sixty-nine tags received fewer votes during round 2 than in round 1, and only three tags received more votes during round 2 (actualities, newsreels, and color films).

Table 4: New Tags for Round 2			
Tags Round 2	**Votes**	**Tags Round 2**	**Votes**
foreign movies	6	publicity stunts	4
origin of films shown	6	regional selections of movies	4
cinema chains	5	seating policy	4
high street / main street / downtown	5	silent film	4
intermittent exhibition (eg only one day a week, only special films)	5	sunday opening	4
map	5	theater chain	4
neighborhood	5	age	3
opening date (broader than opening notice?)	5	censorship data films shown	3
Rural	5	dress code	3
social class	5	ethnicity	3
suburban	5	event	3
audience	4	female exhibitors	3
circuit or chain	4	film entrance	3
comparative studies of popularity and preferences	4	gender	3
downtown	4	imports	3
economic integration	4	photograph	3
film societies	4	postcard	3
geographical expansion	4	press reception of cinema	3
ideological	4	railroad	3
marketing	4	specific films	3
music-hall	4	talkies	3
oral history	4	labor union	2
public hall	4		

Two of the most popular tags added to the collection during round 2 (foreign films, origin of films shown) were concepts that this larger group was able to bring to the list. Just as having a variety of scholars broadened the database schema, having a range of different scholars participating in the tag survey helped to broaden the scope of interested from the original set of tags

While some variation in voting was expected between round one and round two, there were a few terms that were particularly interesting, as there were more than 3 people who changed their minds: children, hour of operation, manager, fair, illustrated song,

organ (a change of 5), circus, men-only screenings, projection problems, and carnival. All of these tags had fewer votes in round 2 than in round 1.

Discussion

This project really had two distinct parts, the terminology, or definitions of the database fields, and the tags. The database fields were the heart of this study and discussed by the participants in more detail. I felt that this project would have benefited from more iterative rounds, as many ideas were introduced in Delphi round one, but there was not an opportunity to fully tease out common meanings within one additional round. I also feel that along with more iterations, this project would also have benefited from even more participants. For example, when the results section was sent back to the scholars for review, Maria A. Velez-Serna commented on the open questions about racial/ political/ ideological association:

> Maybe the difference is that as an 'audience composition' category, racial segregation was based on active, explicit exclusion, whereas ideological, religious or even class-based divisions are less watertight. Perhaps there could be two fields, one being something like 'ethnic/racial composition of audience' (this would cover for instance diasporic audiences, such as in the case of Greek cinemas in Australia) and the other being 'ideological affiliation' (religious, political, clubs, Freemasons, etc). In both of these fields it would be possible to note whether this was an explicit, enforced policy or just a customary one.

While seven was a very respectable number to begin to define complex relationships and definitions, by the end of Delphi round two, there were noticeably different "groups" of scholars who were swaying the conversation – and more varied input might have helped to generalize the discussion. For example while Jeff Klenotic

noted "I think the database covers nearly all the topics, terms, concepts and data of greatest significance to my research." He also mentioned that:

> One concept that I may yet need to tailor into the database has to do with the category of the audience. Two of the central questions of interest in my work are: 1) who chose to attend what cinemas, when, and why? and 2) who chose not to attend what cinemas, when, and why? These are questions relating to variable patterns of cinema attendance and film consumption, and I'm not completely sure how such data will fit into the database diagram.

It is very possible that there are other researchers who grapple with the same issues that Jeff Klenotic feels are still unrepresented by this schema.

The tags were easier to manage, although they were never really defined, so while the group agreed upon a set of tags, there was no discussion about their definitions. One participant noted after completing the second round of tags that it was "somewhat difficult to assess the distinction between my specific current research work and general future research interests." This might be one of the reasons for the variation in voting on some of the selected tags during round 2. There was not as clear of direction on whether the participants should select tags they have used or ones that they might find interesting.

Conclusion and Future Work

This project tried to examine concepts and vocabulary needs of this specific domain of moviegoing as well as the schema, or relationship of these concepts. The ideal outcome of a project such as this would be a complete agreed-upon schema, with very generalized concepts by which all moviegoing scholars can begin their digital

projects, much like the Dublin Core. However when one considers that the Dublin

Core has been a work in progress since 1994 (dublincore.org/about/history/), one

cannot expect to define a field with a few Delphi rounds. I felt that success for this

project would be if the majority of the scholars participating find some benefit to the

sharing of ideas and the iterative process are able to look at the definitions they use

with a broader perspective, and not just in terms of their own research. If the findings

of this project can be used as a reference for moviegoing scholars who begin new

projects, I would also feel that this project was very successful. Maria A. Velez-

Serna, a PhD student at the University of Glasgow who has yet to build a digital

project on her own, but is becoming more involved in this domain said:

> I think the building of the core scheme has been very successful; I can look at
> your diagram now and it gives me an idea of the kind of data that I need to
> gather and how things are connected. I am going to start working in
> organising my data tables into a more structured database soon and I will take
> this as a blueprint. I think the 'tags' part still needs some cleaning, and perhaps
> this would benefit most from input from more people, see if the tags we are
> using are understood and useful for scholars working on different topics too.
> The whole process has been a welcome opportunity to step back and try to
> understand what I mean when I use some words, and most of all it has given
> me some insight into the way that other scholars are working.

The group that ultimately participated in the project was very important in molding

the results. Because there were international scholars as well as scholars from

America, this project was able to uncover some regional/ national concerns that

brought up some very interesting topics for discussion. I do wish that there had been

even more variety in the research which was being done, such as scholars focused on

current-day moviegoing, or moviegoing in China. However, for such a complex

discussion, seven was a good number of people to start with.

While none of the scholars who are currently working on projects came back to say they wanted to change their schemas, they did seem very willing and interested in continuing to tweak the schema definitions in order to "fit" their scholarship and the scholarship of their peers into the schema. Almost every scholar who participated in the process reported that they found the iterative process very interesting, and thought that being forced to think about how other researchers were defining the field was an interesting and worthwhile exercise for them. It seems that a valuable contribution to the discipline to continue this exploration.

While I felt that this project was a great beginning to the development of a common set of terms and elements for scholars within the moviegoing field to work with, there are still many questions left unanswered beyond the "Open Issues" mentioned previously.

Several scholars noted that while this schema is a great "core" to the field of moviegoing, their specific research required additional fields and concepts that were not included in this study. The Dublin Core, which began in 1995 developed a set of 15 basic metadata elements used to describe many different types of physical resources, has basic core of simplified sets of elements. Beginning in 2009 Dublin Core started to develop working groups, or "application profiles" which work to develop specialized schemas that help to address the needs of various sub-communities (http://dublincore.org/). For example, the Dublin Core Education

Schema takes the basic elements of the Dublin Core and adds other elements specific to an educational object, users, duration, learning process, etc. Similarly, this project focused on finding commonalities in order to narrow down the fields to a core set, which would ideally be applicable to all moviegoing projects. Opening up the process to even more moviegoing scholars could provide further refinement of this core set of fields. After this, it would be important to go back and establish working groups to address sub topics that might be relevant to numerous scholars, although perhaps not to the whole community.

Once a common set of definitions are in place and it becomes easier to share data between the various projects, issues of data authority arise. Hired undergraduates or students often fill in the majority of the content for these projects, and while the scholar(s) still has responsibility for the work, they may not review all of it. This noted, what types of citations are needed for these scholars to be able to make use of data prepared by someone unknown? Scholars are often comfortable sharing their data with each other, trusting in their peers to have done quality research. However sharing data with a larger group, when not all of the participants are known, may cause some of these scholars to be skeptical of the data. If this is the case, what types of authority controls will be needed to ensure that the data is still useful?

With the iterative Delphi process that these scholars went through, an effort has begun to define the field of moviegoing in a formalized sense. Hopefully as further collaboration and work between these scholars continues the open discussions will

continue, and these scholars will begin to develop a schema not just with their

individual project in mind, but with the entire spectrum of moviegoing, taking into

account temporal and cultural difference in terminology and research needs.

References

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., et al. (2000). Gene Ontology: tool for the unification of biology. *Nature Genetics*, *25*, 25-29. doi:10.1038/75556.

Ball, A., 2010. Review of the State of the Art of the Digital Curation of Research Data. Project Report. University of Bath, (erim1rep091103ab10).

Barnhart, Benjamin J. (1989). "DOE Human Genome Program". *Human Genome Quarterly* **1**: 1. http://www.ornl.gov/sci/techresources/Human_Genome/publicat/hgn/v1n1/01 doehgp.shtml. Retrieved 2010-03-19.

Baruchson-Arbib, S., and Bronstein, J. (2007). Humanist as Information Users in the Digital Age: the Case of Jewish Studies Scholars in Israel. *Journal of the American Society for Information Science*, *58*(14), 2269-2279.

Brooksbank C, Quackenbush J. Data standards: a call to action. *OMICS 2006*;10:94–9.

Cantor, C. R. (1990). Orchestrating the Human Genome Project. *Science*, *24*(8), 49-51.

Ciffolilli, A. (2003). Phantom authority, self–selective recruitment and retention of members in virtual communities: The case of Wikipedia. *First Monday*, *8*(12).

Chang, R., Kennedy, G., & Petrovic, T. (2008). Web 2.0 and user-created content: Students negotiating shifts in academic authority. *Proceedings ascilite Melbourne 2008*, Australia, 165-169.

Chesney, T. (2006). An empirical examination of Wikipedia's credibility. *First Monday*, *11*(11).

Cook-Deegan R (1989). "The Alta Summit, December 1984". *Genomics* 5 (3): 661–3. doi:10.1016/0888-7543(89)90042-6. PMID 2613249.

Davis, Glyn, (2007, October 31). The end of the gatekeeper. *The Age*. Retrieved from http://www.theage.com.au/

Dibbets, Karel. (n.d.). About. In *Cinema Context Films*. Retrieved May 06, 2010, from http://www.cinemacontext.nl.

Frenkel, K. A. (1991). The Human Genome project and informatics: a monumental scientific adventure. *Communications of the ACM*, *34*(11), 40-51.

Garcia, J., and Steinmueller, W. E. (2003, June). Applying the Open Source Development Model to Knowledge Work. Science and Technology Policy Research.

Green, Rebecca. (2000). Locating Sources in Humanities Scholarship: The Efficacy of following Bibliographic References, *The Library Quarterly*, 70(2), 201-229.

Harter, S. (1997). Scholarly Communication and the Digital Library: Problems and Issues. Journal of Digital Information, 1(1). Retrieved from http://journals.tdl.org/jodi/rt/printerFriendly/4/4.

Human Genome Program (1993) *Human Genome News* 4 (5).

Nicholson, D. P. J. (2006). The Changing Role of Intellectual Authority. Notes for remarks to the Association of Research Libraries 148th Membership Meeting, Canada.

Lacher, Martin S., and Greorg Groh. (2001). Facilitating the exchange of explicit knowledge through ontology mappings, *FLAIRS-01 Proceedings*, American Association for Artificial Intelligence.

Ovadia, S. (2007). Digg.com and Socially-Driven Authority, Steven Ovadia. *Library Philosophy and Practice 2007*, Retrieved from http://digitalcommons.unl.edu/libphilprac/.

Regalado, M. (2007). Research Authority in the Age of Google: Equilibrium Sought, Mariana Regalado. *Library Philosophy and Practice,* Retrieved from http://digitalcommons.unl.edu/libphilprac/.

Rosenberg, L. C. (1991). Update on National Science Foundation funding of the "collaboratory", *Communications of the ACM, 34*(12), 83, New York: ACM Press.

Rosenthal, et al. (2004). From Semantic Integration to Semantics Management: Case Studies and a Way Forward, SIGMOD Record, 33(4), 44-50.

Seaman, D. (2004). Aggregation, Integration, and Openness: Current Trends in Digital Libraries. *Report from the International Symposium on Digital Libraries and Knowledge Communities in Networked Information Society,* (DLKC'04).

Staab, et. Al. (2000). Semantic community Web portals, *Computer Networks*, 33(1-6), 473-491.

Tabb, K. (2008). Authority and Authorship in a 21st-Century Encyclopaedia and a 'Very Mysterious Foundation'. *eSharp*, (12), 1-21.

Watson, James D. (1990). The Human Genome Project: Past, Present, and Future. *Science*, *24*(8), 44-49.

Appendices

Appendix 1: "Going to the Show" Experts Meeting

"Going to the Show" digital collection
Experts meeting on October 15, 2008

Blurbs about Oct 15

We have assembled a group of Scholars who are researching the relationship of film history and the social experience of cinema. All of them are trying to use the new digital technologies to help re-imagine the history of cinema, by expanding the scope beyond the study of individual films taken out of their historical context, and placing the experience of cinema within historical context. Each of the panel members have their own background experience working on similar digital/ film projects.

We are going to benefit from their reviewing our project at it's current state of development, by helping us to continue to think through how we can best display, store, and organize our data. And how our project might be of use to a wide range of potential customers. Each of these scholars has connections to other fields, bringing these perspective to bear as well: Social History, American Studies, History of Race Relations. They will also help us to think about the comparative possibilities of this project. The geographic focus of their work ranges from the South Eastern United states to New England, to Australia.

Attendees for Oct 15

Richard Maltby
Richard Maltby is Professor of Screen Studies. He moved to Flinders from the UK, where he established the Bill Douglas Centre for the History of Cinema and Popular Culture at the University of Exeter, before becoming Research Professor in Film Studies at Sheffield Hallam University. He is the author of Hollywood Cinema and Harmless Entertainment: Hollywood and the Ideology of Consensus, as well as numerous articles on American cinema and popular culture. He is currently working on a history of regulation and the politics of Hollywood in the 1920s and 1930s, to be called Reforming the Movies: the Governance of the American Cinema, 1908-1939.
http://ehlt.flinders.edu.au/screen/staff/maltby.php

Arthur Knight
Arthur Knight, who works on film, mass and popular culture, and 20th century literature, received his PhD from the University of Chicago (Department of English) in 1998. He has been teaching at the College from 1993. He has co-edited the collection Soundtrack Available: Essays on Film and Popular Music (Duke UP, 2001) and written Disintegrating the Musical: Black Performance and American Musical Film (Duke UP, 2002) as well as on jazz in film and African American conceptions of stardom. Currently he is at work on several projects: an essay on "whiteface"; an essay on Spike Lee; a book-length study of African Americans and stardom; and several collaborative local history projects.
http://www.wm.edu/americanstudies/research.php?personid=77980
http://www.wm.edu/americanstudies/wtp/
http://www.homerproject.org/?display=icarg

Jeff Klenotic
Jeffrey Klenotic is an Associate Professor of Communication Arts at the University of New Hampshire at Manchester where he teaches media history, media literacy and cultural studies. In 1997 he received the UNH Manchester Excellence in Teaching Award.
Dr. Klenotic maintains a research program focusing on media and cultural history. His work appears in academic journals and books, including Film History: An International Journal, The Communication Review, Velvet Light Trap, The Sounds of Early Cinema, Going to the Movies, The New Cinema History, The Encyclopedia of Film, and The Encyclopedia of Early Cinema.
http://www.jeffklenotic.com/

Rob Nelson

Rob Nelson is an historian of the nineteenth-century United States and a technologist. For eight years he was the Technical Liaison to the Humanities at the College of William & Mary. Starting in 2008 he became the Associate Director of the Digital Scholarship Lab at the University of Richmond. He has worked on several digital humanities projects, including the Williamsburg Theater Project
http://oncampus.richmond.edu/news/oct07/digital.html
http://web.wm.edu/americanstudies/wtp/?&=&svr=www

James Burns
Professor Burns, the 2005 winner of the Gentry Award for outstanding teaching, has came to Clemson in 1999. He is a specialist in African history, though he also teaches courses in Film and History, Historical Methods, and Western Civilization. Dr. Burns's first book, Flickering Shadows: Cinema and Identity in Colonial Zimbabwe, a study of how European leaders used film as an imperial tool and how colonial Africans resisted in unexpected and surprising ways, was named by Choice magazine as one of its Outstanding Academic Titles for 2002. He is the co-author with Robert O. Collins of A History of Sub-Saharan Africa, which was published by Cambridge University Press in 2007. He is currently working on a book about the early history of the cinema throughout the African Diaspora.
http://www.clemson.edu/caah/history/faculty/burns.htm

Appendix 2: "Going to the Show" Experts Meeting Schedule

<u>**Tentative Schedule of Oct 15**</u>

8:45 Hello

9am
Review the Going to the Show Goals and Timeline (Bobby and Natasha)
maybe have Bobby talk a little bit and some about getting funding for such a project.
Talk about Connection to Doc South (Natasha)

10:00
Bathroom

10:15
Visitors present their projects.

11:45
Break

12:00
Tour of North Carolina Collection (Nick - Jason..)

Tour of Digital Production Center (Fred)

1 pm
LUNCH *have lunch here by 12:30*

1:45
Discuss the Process Details. (Kevin)
software choices..
why google maps,

2:15
Review the database structure (Cliff, Kevin, Hugh)
Some of the topics we discussed - venue, date issue, etc.

3:00
BREAK

3:15
Interface Review.
This is where we would really like their input - with talks about the advanced search, and how we physically show the connections between content and theaters, etc.

4:30
Closing Discussion - possibilities for collaboration - how to move forward.
debreifing - open up the discussion about how to progress into the future.

5:30
To Franklin St for drinks and collaboration.

Appendix 3: Participants of the Delphi Study

Jeffrey Klenotic
Jeffrey Klenotic is an Associate Professor of Communication Arts at the University of New Hampshire at Manchester where he teaches media history, media literacy and cultural studies. In 1997 he received the UNH Manchester Excellence in Teaching Award.

Dr. Klenotic maintains a research program focusing on media and cultural history. His work appears in academic journals and books, including Film History: An International Journal, The Communication Review, Velvet Light Trap, The Sounds of Early Cinema, Going to the Movies, The New Cinema History, The Encyclopedia of Film, and The Encyclopedia of Early Cinema.
http://www.jeffklenotic.com/

Daniel Biltereyst
Professor Daniel Biltereyst
Chair - Dept. of Communication Studies
Daniel Biltereyst is Professor in Film, Television and Cultural Media Studies at the Gent University, Belgium, where he leads the Centre for Cinema and Media studies (CIMS). His work is on film audiences, historical reception and cinemagoing, as well as on controversial film, television and censorship, situated within moral panic and public sphere theories.

Karel Dibbets
Universiteit van Amsterdam
Media Studies
Karel Dibbets is Lecturer in Media History at the Universiteit van Amsterdam. He is an expert in the history of cinema culture in the Netherlands, and he has a keen interest in the development of digital knowledge infrastructures. He is the architect and editor of the website annex research instrument **Cinema Context**, which was funded by the Netherlands Organisation for Scientific Research (NWO).
http://www.cinemacontext.nl/

Philippe Meers
Philippe Meers is an associate professor in film and media studies in the Department of Communication Studies at the University of Antwerp, where he is research leader in the Visual Culture Research Group. Recent publications in Watching The Lord of the Rings (Lang, 2007); Illuminace (2008); Film – Cinema – Spectator: Film Reception (Schüren 2010) The Handbook of Political Economy of Communications (Blackwell 2011). He was promoter of "the Enlightened City" project (2005-2008) and is currently (co-)promoter of research projects on 'Antwerp Cinema City', 'Cinema and diaspora' and 'Cinema culture in Monterrey, Mexico'. With R. Maltby and D. Bilterest, he is editing The New Cinema History (Wiley-Blackwell, 2010) and Audiences, Cinema and Modernity (forthcoming). A research note (together with Daniel Biltereyst and Lies Van De Vijver) on "the Enlightened City" is forthcoming in Screen, 51/3 Fall 2010.
www.ua.ac.be/philippe.meers

Arthur Knight
Arthur Knight, who works on film, mass and popular culture, and 20th century literature, received his PhD from the University of Chicago (Department of English) in 1998. He has been teaching at the College of William & Mary from 1993. He has co-edited the collection Soundtrack Available: Essays on Film and Popular Music (Duke UP, 2001) and written Disintegrating the Musical: Black Performance and American Musical Film (Duke UP, 2002) as well as on jazz in film and African American conceptions of stardom. Currently he is at work on several projects: an essay on "whiteface"; an essay on Spike Lee; a book-length study of African Americans and stardom; and several collaborative local history projects.
http://www.wm.edu/americanstudies/research.php?personid=77980
http://www.wm.edu/americanstudies/wtp/

http://www.homerproject.org/?display=icarg

Robert Allen
Robert Allen's teaching and research interests are broad and interdisciplinary. His research focuses on American cultural history (particularly the history of American popular entertainment) and the application of digital technologies to history. He has written on the history of U.S. radio and television (Speaking of Soap Operas, 1985), film history and historiography (Film History: Theory and Practice, 1985), and American popular theater of the nineteenth and early twentieth century (Horrible Prettiness: Burlesque and American Culture, 1992). He is also the editor of To Be Continued: Soap Operas Around the World (1995) and of two editions of the widely-used television criticism anthology, Channels of Discourse and Channels of Discourse, Reassembled (1987, 1994). He is the co-editor of The Television Studies Reader (2004) and Going to the Movies: Hollywood and the Social Experience of Cinema (2007).

"Going to the Show" (www.docsouth.unc.edu/gtts), an online digital library documenting the history of moviegoing in North Carolina, was launched in 2009. Professor Allen was awarded one of the first National Endowment for the Humanities Digital Humanities Fellowships in 2008-09 for his work on this project, which was also honored as a "We the People" project by the NEH.

He is currently working on another digital humanities project in collaboration with UNC's Wilson Library. "Main Street, Carolina" will provide cultural heritage organizations in 45 towns and cities across the state with a digital plat form for creating and managing digital content illuminating the history of their downtowns and displayed on historic maps. For this project Professor Allen received the first C. Felix Harvey Award to Advance Institutional Priorities at UNC-CH in 2009.   His teaching interests include the history o f American film and media, globalization and national identity, the family and social change in America, digital history, and comparative social and cultural history (especially American and Australian histories).

John Caughie
Director, Arts Lab, Faculty of Arts, University of Glasgow

John Caughie is Professor of Film & Television Studies at the University of Glasgow. From 1999 to 2005, he was Dean of the Faculty of Arts, a Faculty which contains thirteen departments in the fields of arts and humanities. He is an active researcher in the field of film and television with current interests in film theory, television drama, and screen adaptations of Charles Dickens. He is an editor of Screen, a leading international journal in film and television studies published by Oxford University Press. He was co-editor with Charlotte Brunsdon of the series Oxford Television Studies which has published twelve volumes in the last six years. He convenes the Scottish Funding Council's working group on Knowledge Transfer and Cultural Engagement.   He currently serves on the Council of the Arts and Humanities Research Council, the national Research Council responsible for funding research and postgraduate study in the Arts and Humanities. He chairs its Research Committee, its programme on ICT, and has particular responsibilities for a number of its Research Centre.

Maria A. Velez-Serna
PhD student
Research in early film distribution in Scotland
University of Glasgow

Appendix 4: Cinema Context Entities

CCC Entities and fields

Film

Title
Year
Country
Directed by
Production Co.
Length
Smalfilm (Y/N)
Short/Long
IMDBID
Info

Censorship

Institute
File number
Date
Rating
Rating info
Length
info

Cinema

Name
Info

Programme

Name
Info

Accom. Info
Book keeping

Person

Last name
Other names
First names
Middle
Suffix
Date of birth
City of birth
Date of death
Place of death
Biography
Info

Company

Name
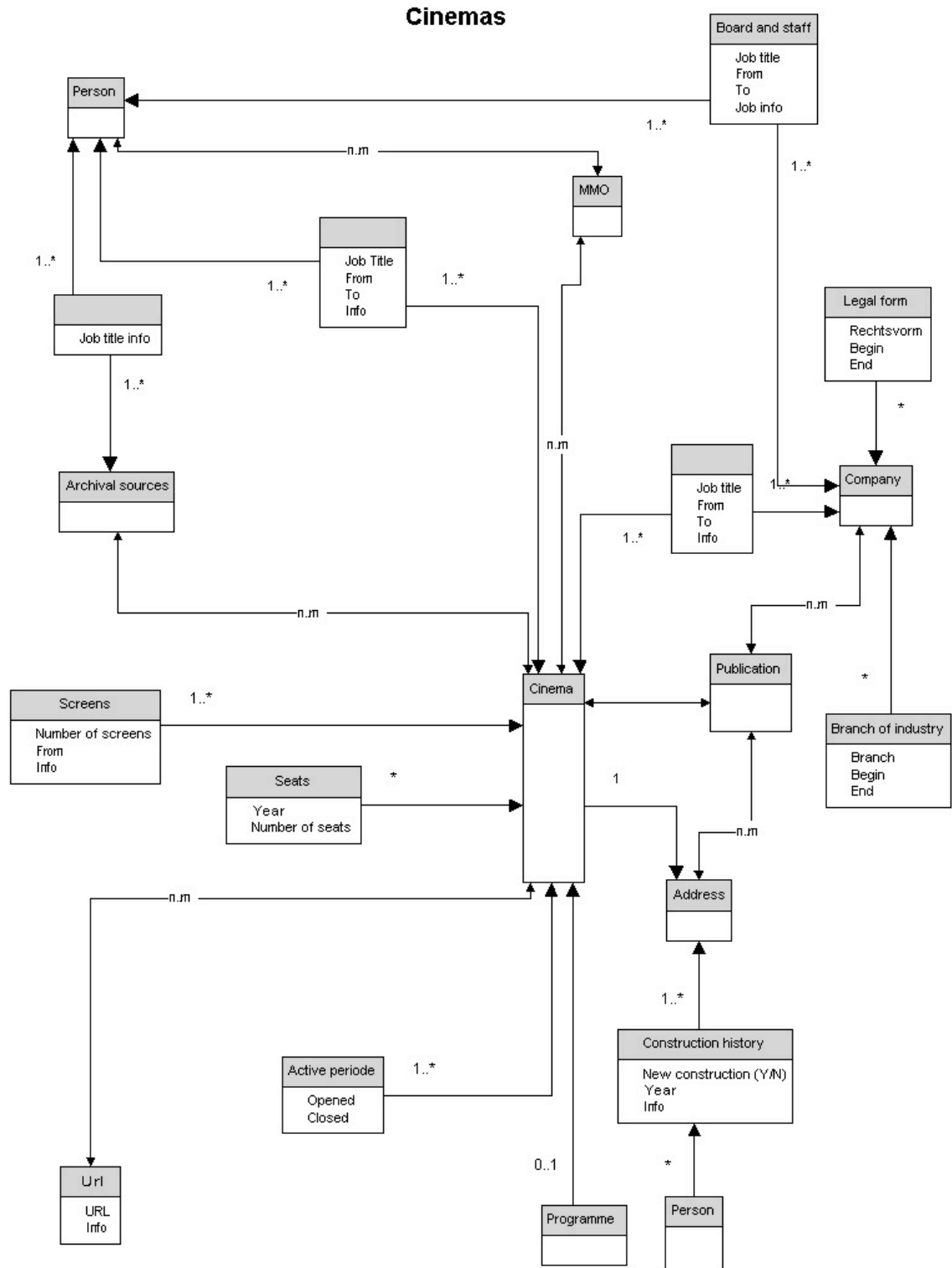Other names
Date established
Date disbanded
Info

Address

XY
Street
Former street name
City
Former city name
Info

Publication

Article/Chapter
Periodical
Authors
Editors
Date periodical
Year published
Volume
Number
Place
Publisher
Pages
Type
Info
ISBN/ISSN

MMO

Type
category
Title
Description
Copyright/Source
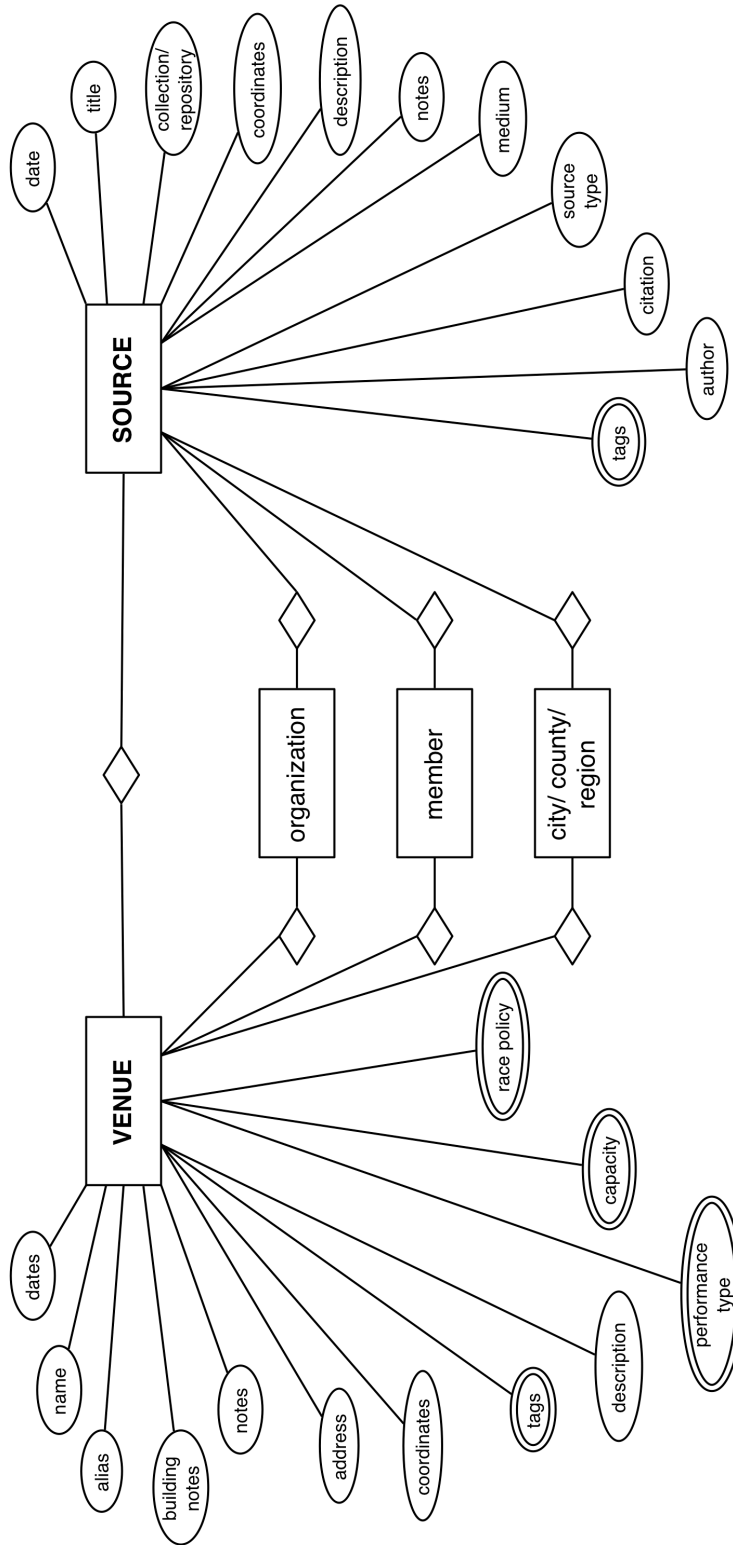Digitally available (y/n)
ID (external)
Url

Archival sources

Institute
Collection
File number
Descr.
Info
Digitally available (y/n)
ID (External)

Appendix 5: Cinema Context Schema

Cinemas

Board and staff
- Job title
- From
- To
- Job info

Person

MMO

Job Title
- From
- To
- Info

Job title info

1..*

1..*

n.m

1..*

1..*

Legal form
- Rechtsvorm
- Begin
- End

*

Archival sources

Job title
- From
- To
- Info

Company

1..*

n.m

1..*

n.m

Screens
- Number of screens
- From
- Info

1..*

Cinema

Publication

Branch of industry
- Branch
- Begin
- End

*

Seats
- Year
- Number of seats

*

1

Address

n.m

n.m

1..*

Construction history
- New construction (Y/N)
- Year
- Info

Active periode
- Opened
- Closed

1..*

Url
- URL
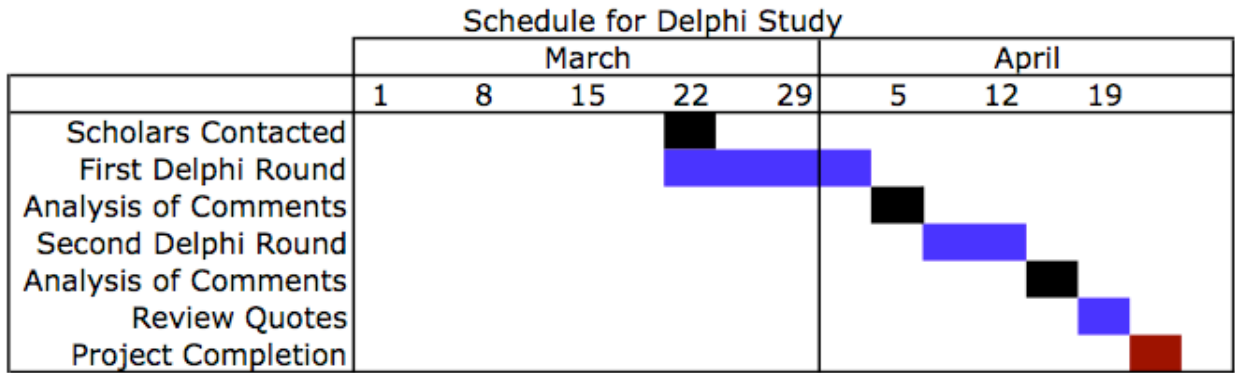- Info

0..1

Programme

*

Person

Appendix 7: "Going to the Show" Abstracted Schema

Going to the Show

this is an abstracted representation of our actual er diagram

Appendix 8: Projected Delphi Schedule

			Schedule for Delphi Study					
		March				April		
	1	8	15	22	29	5	12	19
Scholars Contacted				■				
First Delphi Round					▬			
Analysis of Comments						■		
Second Delphi Round							▬	
Analysis of Comments							■	
Review Quotes								▬
Project Completion								■

Appendix 9: Initial Combined Schema (Delphi Round 1)
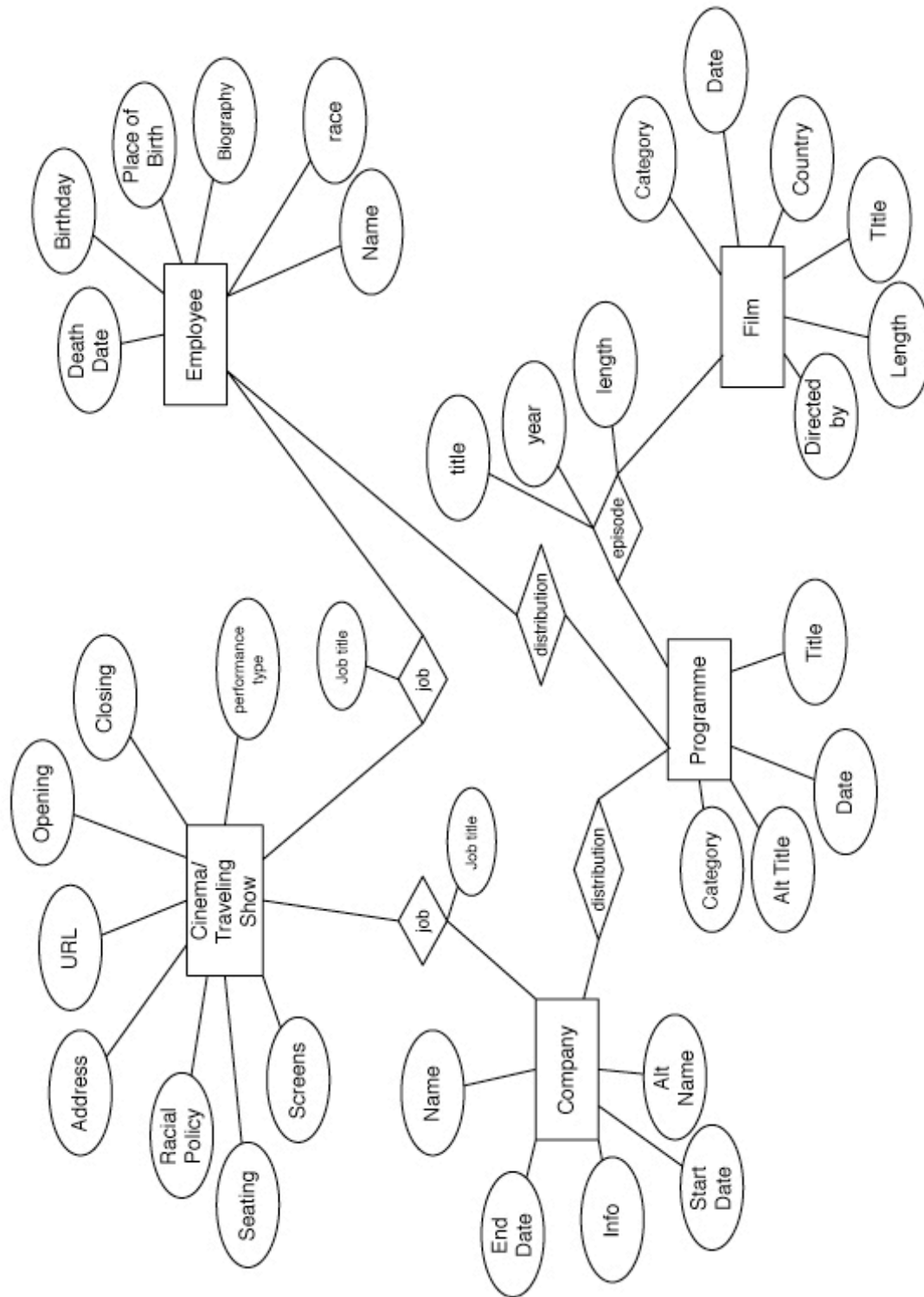
Appendix 10: Initial Data Dictionary (Delphi Round 1)

Instructions

This is the "meat" of the study. It is an attempt to try and define concepts, which are structured with tables and fields of a database.

For each field, please mark if it is an essential bit of information for your research or not – for example, if you are interested in studying when movie theaters moved from small operations to large networks run by major picture studios, then you would most likely find the following fields useful for your research: Cinema Seating, Cinema Screen, Cinema Opening Date, Cinema Closing Date, and perhaps Company Name, and Company Job.

Next you are asked: "Have you ever used this info in your research, and if so how. If not, are you likely to? If so how?" This is where you are given the opportunity to describe exactly what type of data you would be looking to enter into this field. If you have used this type of information in your research, a few examples would be extremely helpful. If you have not, please write a sentence or two about what you might use this type of information for.

Throughout the list you will notice a few instances where I mention a "controlled vocabulary". Controlled vocabularies are very important tools within databases, as they provide a pre-defined way of representing information, and they allow simple grouping of like things. They do, however limit your ability to specifically describe a bit of information. One example I have been struggling with is Cinema size. If this is listed as Small, Medium, or Large, then this would allow a researcher or visitor to easily navigate through different sizes of cinemas and determine patterns related to cinema size. However, this requires defining Small, Medium and Large, which may vary for different locations. Please indicate if you agree a controlled vocabulary fits the noted field, and if so, what types of values would be useful.

At the end of each "table" you will be asked to describe other types of information about this particular concept you would like to collect. This is where you are able to discuss any bit of information that was not covered already. This may be written in paragraph form, or as a list – with a brief description of what each bit of information means to you.

Again, if any part of this is confusing, please note the confusing portion on this form and continue as best you can, and I will work to try and explain to the best of my ability so that you are able to respond with informed thought.

Thanks.

Cinema/ Traveling Show

Screens   Number of screens in the cinema
- Is this an essential bit of information for your research?   YES   NO
- Have you ever used this info in your research, and if so how. If not, are you likely to? If so how?

Seating Amount of Seating available
             In numbers, or S, M, L?
- Is this an essential bit of information for your research?   YES   NO
- Have you ever used this info in your research, and if so how. If not, are you likely to? If so how?

Racial Policy   Needs controlled vocab
                 (black, white, mixed seating, segregated balcony, segregated
      showings, …)
- Is this an essential bit of information for your research?   YES   NO
- Have you ever used this info in your research, and if so how. If not, are you likely to? If so how?

Address   Address of the Cinema
- Is this an essential bit of information for your research?   YES   NO
- Have you ever used this info in your research, and if so how. If not, are you likely to? If so how?

URL   If there is a current URL for the Cinema
- Is this an essential bit of information for your research?   YES   NO
- Have you ever used this info in your research, and if so how. If not, are you likely to? If so how?

Opening   Date of opening in particular location
- Is this an essential bit of information for your research?   YES   NO
- Have you ever used this info in your research, and if so how. If not, are you likely to? If so how?

Closing Date of closing in particular location
- Is this an essential bit of information for your research?   YES   NO
- Have you ever used this info in your research, and if so how. If not, are you likely to? If so how?

Performance Type   Types of performances
- Is this an essential bit of information for your research?   YES   NO
- Have you ever used this info in your research, and if so how. If not, are you likely to? If so how?

OTHER NOTES ABOUT CINEMA/ TRAVELLING SHOW
- Please describe what other types of information about Cinemas or Travelling Shows you would like to collected or study.

<u>Company</u>

Name
- Is this an essential bit of information for your research?   YES   NO
- Have you ever used this info in your research, and if so how. If not, are you likely to? If so how?

Alternate Name
- Is this an essential bit of information for your research?   YES   NO
- Have you ever used this info in your research, and if so how. If not, are you likely to? If so how?

Start Date   First recorded existence of company
- Is this an essential bit of information for your research?   YES   NO
- Have you ever used this info in your research, and if so how. If not, are you likely to? If so how?

End Date   Last recorded existence of company
- Is this an essential bit of information for your research?   YES   NO
- Have you ever used this info in your research, and if so how. If not, are you likely to? If so how?

Info   A long text field to allow other descriptive information about the company
- Is this an essential bit of information for your research?   YES   NO
- Have you ever used this info in your research, and if so how. If not, are you likely to? If so how?

Job   The Company's relationship with the Theater
          (supplier, production, financier, etc – need controlled vocabulary)
- Is this an essential bit of information for your research?   YES   NO
- Have you ever used this info in your research, and if so how. If not, are you likely to? If so how?

OTHER NOTES ABOUT COMPANY
- Please describe what other types of information about Companies you would like to collected or study.

Employee
An employee of the Cinema or traveling show

Name
- Is this an essential bit of information for your research?   YES   NO
- Have you ever used this info in your research, and if so how. If not, are you likely to? If so how?

Birthday   YYYY-MM-DD, YYYY-MM, YYYY
          Or any amount of data known.
- Is this an essential bit of information for your research?   YES   NO
- Have you ever used this info in your research, and if so how. If not, are you likely to? If so how?

Place of Birth City, State, or Country
- Is this an essential bit of information for your research?   YES   NO
- Have you ever used this info in your research, and if so how. If not, are you likely to? If so how?

Death Date   YYYY-MM-DD, YYYY-MM, YYYY
          Or any amount of data known.
- Is this an essential bit of information for your research?   YES   NO
- Have you ever used this info in your research, and if so how. If not, are you likely to? If so how?

Race   Need controlled vocabulary
- Is this an essential bit of information for your research?   YES   NO
- Have you ever used this info in your research, and if so how. If not, are you likely to? If so how?

Biography   A long text field to allow other descriptive information about the person.
- Is this an essential bit of information for your research?   YES   NO
- Have you ever used this info in your research, and if so how. If not, are you likely to? If so how?

Job   The job title
- Is this an essential bit of information for your research?   YES   NO
- Have you ever used this info in your research, and if so how. If not, are you likely to? If so how?

OTHER NOTES ABOUT EMPLOYEE
- Please describe what other types of information about Employees you would like to collected or study.

<u>Programme</u>
A collection of shows for preview in a specific Cinema

Title
- Is this an essential bit of information for your research?   YES   NO
- Have you ever used this info in your research, and if so how. If not, are you likely to? If so how?

Alt. Title
- Is this an essential bit of information for your research?   YES   NO
- Have you ever used this info in your research, and if so how. If not, are you likely to? If so how?

Date   Date of this specific Program
        YYYY-MM-DD, YYYY-MM, YYYY Or any amount of data known.
- Is this an essential bit of information for your research?   YES   NO
- Have you ever used this info in your research, and if so how. If not, are you likely to? If so how?

Category   Need Controlled Vocabulary for this
            Illustrated Song, Mixed Program, Horror, etc)
- Is this an essential bit of information for your research?   YES   NO
- Have you ever used this info in your research, and if so how. If not, are you likely to? If so how?

Distribution Company The Company that distributes the Programme
- Is this an essential bit of information for your research?   YES   NO
- Have you ever used this info in your research, and if so how. If not, are you likely to? If so how?

Distribution Employee A person that distributes the Programme
- Is this an essential bit of information for your research?   YES   NO
- Have you ever used this info in your research, and if so how. If not, are you likely to? If so how?

OTHER NOTES ABOUT PROGRAMME
- Please describe what other types of information about Programmes you would like to collected or study.

Film

Title
- Is this an essential bit of information for your research?   YES   NO
- Have you ever used this info in your research, and if so how. If not, are you likely to? If so how?

Date   YYYY-MM-DD, YYYY-MM, YYYY
     Or any amount of data known.
- Is this an essential bit of information for your research?   YES   NO
- Have you ever used this info in your research, and if so how. If not, are you likely to? If so how?

Country
- Is this an essential bit of information for your research?   YES   NO
- Have you ever used this info in your research, and if so how. If not, are you likely to? If so how?

Directed by
- Is this an essential bit of information for your research?   YES   NO
- Have you ever used this info in your research, and if so how. If not, are you likely to? If so how?

Length
- Is this an essential bit of information for your research?   YES   NO
- Have you ever used this info in your research, and if so how. If not, are you likely to? If so how?

Category
- Is this an essential bit of information for your research?   YES   NO
- Have you ever used this info in your research, and if so how. If not, are you likely to? If so how?

OTHER NOTES ABOUT FILM
- Please describe what other types of information about Films you would like to collected or study.

Film Episode

Title
- Is this an essential bit of information for your research?   YES   NO
- Have you ever used this info in your research, and if so how. If not, are you likely to? If so how?

Year
- Is this an essential bit of information for your research?   YES   NO
- Have you ever used this info in your research, and if so how. If not, are you likely to? If so how?

Length
- Is this an essential bit of information for your research?   YES   NO
- Have you ever used this info in your research, and if so how. If not, are you likely to? If so how?

OTHER NOTES ABOUT FILM EPISODE
- Please describe what other types of information about Film Episodes you would like to collected or study.

Appendix 11: Initial Tag Survey (Delphi Round 1)

1. Introduction.
These tags are an effort to create a controlled vocabulary of common tags, themes, or subjects which are of interest to scholars in the field. The hope is to come up with a core set of tags, which could be used across projects in an effort to make data more interoperable. It is understood that each project will have it's own unique focus and set of descriptors which might not be relevant across collections.

2. Tags
 Please select any of the terms you would find relevant to your research.
1. Please select all that would apply to your research/ interests

- WWI
- actualities
- admission price
- advertisement
- air conditioning
- airdome
- alias
- american indians
- amusement park
- architect
- balcony
- balcony, racially segregated
- band or orchestra
- birth of a nation
- black theater
- boxing films
- carnival
- censorship
- change in owner
- charity
- children
- church
- circus
- closing
- color films
- company
- concessions
- contest
- crime
- distribution
- drugstore

- fair
- family
- fashion show
- feature film
- fight films
- film title
- fire
- hotel
- hours of operation
- illustrated song
- influenza
- ladies and children
- lecturer
- licensing
- local films
- location change
- manager
- men only screenings
- midnight shows
- minstrel shows
- morality
- motion picture party
- music
- name change
- newsreels
- non-theatrical venue
- notable
- open-air screening
- opening notice
- opera house
- organ

- penny arcade
- photographer
- pianist
- program change
- projection problems
- projectionist
- promotions
- proposed theater
- public appearance
- purpose-built movie theater
- racial policy
- re-opening
- religious films
- religious objection
- renovation
- school
- seasonal exhibition
- seating capacity
- separate entrance
- serials
- social club
- sound films
- special showing
- storefront theaters
- streetcars
- tent
- traveling exhibitor
- vaudeville
- ventilation
- women only screening
- ymca

2. Please list any topics you feel should be included in this list. This may also included broader or more narrow terms if you feel those would be more effective.

Appendix 12: Combined Schema (Delphi Round 2)
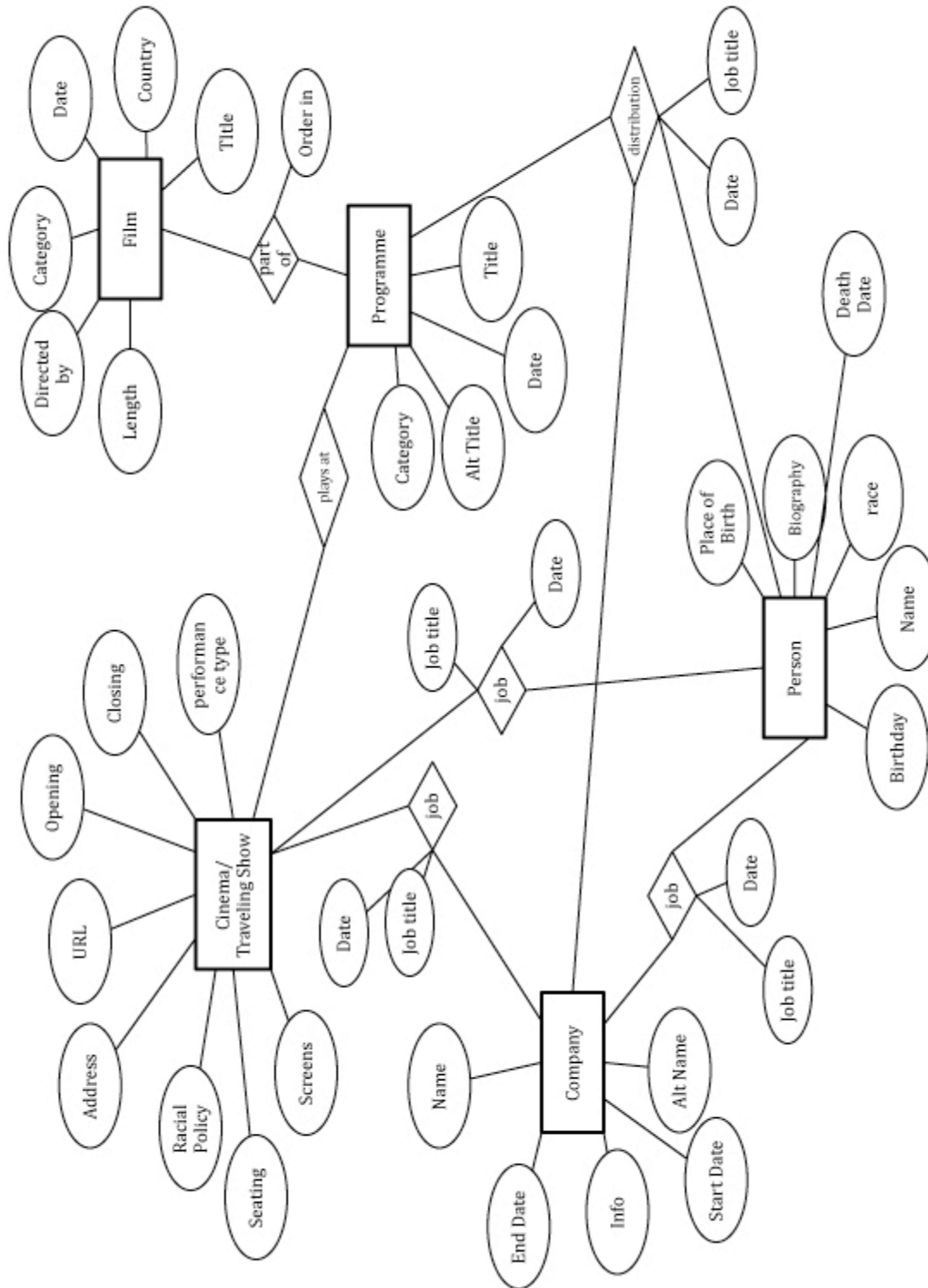
Appendix 13: Data Dictionary (Delphi Round 2)

<u>**Instructions**</u>

I have gone through the responses of the first part of the Delphi Study, and tried to integrate any new ideas, as well as bring up any discrepancies between definitions of fields, that I saw. I have included quotes and descriptions that might help to shape the definition of these fields.

I have colored red, any new fields, as well as any questions about previous fields, which might help to weed out a shared definition. Please be sure to comment on all of the sections colored red, and feel free to read through the other comments and write notes or comments about things that you may agree with or disagree with.

Please use the "Track Changes" tool within Microsoft Word, **<u>or</u>** type your responses in another color other red – so I can easily see where you have commented.

Again, please don't hesitate to email me with any questions as you are going through this.

Thanks.

<u>**Cinema**</u>

"It is necessary to differentiate between two entities: address and cinema. Running a cinema is not the same as owning or developing a building. A venue can change its management, its name and/or its function without altering the property conditions – and vice versa.
Our database uses the address document to describe the property ownership, the architect, and major stages of the building history - information that is usually outside the realm of cinema history. On the other hand, the cinema document in our database describes seating capacity, number of screens, staff, dates of opening and closing, etc., and it has a link to the address document."

"Travelling show
A travelling show is an event (or a company), while a cinema is a building. A travelling show can happen anywhere, outside or inside a building, but the building is not a travelling show. A travelling show does not change the function of a building, e.g. a town hall. A cinema remains a cinema, even when it is closed after a fire."

How do you define a cinema? Is it sufficient to make a "travelling show" a company – with the attributes of a company and not a cinema?

Cinema name
- Is this an essential bit of information for your research?   YES NO
- Have you ever used this info in your research, and if so how. If not, are you likely to? If so how?

Screens
Number of screens in the cinema + Date (there could be more than one pairing per Cinema-allowing to show the shift in number of screens within a single venue.)

> "The number of screens in important in studying the changing nature of moviegoing in the 1970s and 80s."

Cinema Seating Capacity + Date   Amount of Seating available
SML**,** seems to be too relative of a measure – perhaps a scale. 0 – 150, 150 – 300, > 300 ?
Or precise number if available.

> "In general, again, it depends on the quality of your data (cf. do you have very reliable data on the precise number of seats?) whether you can work with precise numbers."

> "I could see S/M/L working, since often it's hard to trace exact seat counts (and they can change in the record somewhat randomly). I think perhaps more than three categories would be good (five? six?) and providing specific number ranges (0-250, say, etc.) would be very important."

> "S, M, L are likely to be relative to urban or rural location – i.e. most rural cinemas would be Small relative to urban, but some may be Large relative to other rural. It would therefore be necessary to use seating capacity numbers"

Racial Policy

This does not seem to be a consistent factor, however is extremely important in studies where it is used. I'm leaving it in with the note that it is non-essential for sharing of information.

"My research has been on African American filmgoing, so this is crucial for me."

Address   Address of the Cinema
Address seems to be a very important field for everyone. Please mark to what level of detail you need to be useful –

GIS coordinates,
street address
district
city
county
state
country.

~~URL~~
There is not enough interest in this field to include it…

Opening/ Closing   Date of opening & closing in particular location
The opening date of a cinema is essential for everyone – but the consensus is that the closing date is harder to verify.
Would the closing date still be useful if it were estimated (est. closing date: 1966)?

> **"**I have not used closing date because this info is very hard to find and verify. If I had it, however, I would definitely use it, in particular to be able to quantify and map existing cinemas at different moments. It would provide much more accurate reports from database queries and better map animations."

Performance Type   Types of performances
Is this field necessary if "programs" denote whether they are "mixed programs" or not?

> "I have only touched on this in my research, but I think it's crucial to understand the cinema as an (often) "mixed" venue, and to be able to track how that changes over time."

> "Key data for charting the movement from music hall or fairground to fixed-site cinema."
> This might be covered if the "building type" field were used – would this be sufficient?

> "I make notes on whether it was a continuous show, or if not, whether there were one or two shows per night. I also make a note of whether it included vaudeville turns or other attractions, eg phonograph records, illustrated songs, elocutionists."

> "We think that a function or type will become clear by looking at what happens in the venue. Wherever a movie is shown, it counts as a venue, and becomes part of cinema history, even when it is not seen as a cinema proper. 'Performance type' is always temporary and it needs dates in case of type changes. In my opinion, it might be sufficient to mention a performance type in a general note describing the venue."

Ticket price range + Date
- Is this an essential bit of information for your research?   YES NO
- Have you ever used this info in your research, and if so how. If not, are you likely to? If so how?

Building type

"What the hall was before being a cinema (a church, a public hall, the Salvation Army, a skating rink, a shop, a tea-room, a music hall, etc); if it was purpose-built, the name of the architect or building contractor."
- Is this an essential bit of information for your research?   YES NO
- Have you ever used this info in your research, and if so how. If not, are you likely to? If so how?

Ideological Association
- Is this an essential bit of information for your research?   YES NO
- Have you ever used this info in your research, and if so how. If not, are you likely to? If so how?

Technology + Date

"A cinema changes its technical installation several times during its history. I have used the dates of its conversion from silent to sound cinema to plot the geographical diffusion of an innovation. Others may want to focus on the introduction of wide screen, stereo, etc. It would be enough, I think, to add a date with some information."
- Is this an essential bit of information for your research?   YES NO
- Have you ever used this info in your research, and if so how. If not, are you likely to? If so how?

OTHER NOTES ABOUT CINEMA/ TRAVELLING SHOW

"What is crucial is the number of people attending/paying the shows. We had some data for very specific cinemas for specific periods."
Would this not be covered within the size of the cinema? Or fif you often have data for the number of people actually attending a film?

"Distribution chains and networks."
This information would be included in the Company – Job – Cinema relationship

"Local authority licensing."
I don't understand this - please clarify..

Circuit: Whether cinema belonged to a chain or was operated in conjunction with at least one other venue, especially for film-booking purposes.
This information would be included in the Company – Job – Cinema relationship

Owner/proprietor/lessee (who was responsible for the actual building) and manager (who ran the show); manager should have start and end date, to be able to account for changes of management.
This information would be included in the Person – Job – Cinema relationship

Data on reception, audiences, marketing, oral history data on experiences, social class and cinema going conventions

Source material for films, cinemas, and companies will hopefully exist for many of your databases. I have opted to leave this out of the schema. I do think that they are important, and should be included whenever possible, but when merging databases, I think it's best to deal with just the "data", and allow the sources, to be browsed, or exchanged through a separate process.

Link with genres/origin of films screened in the theatre.
This would be covered within the relationship of a particular film to a cinema.

Sources and publications
Our database has a facility to add information about sources and publications about a venue or other major entities. We don't collect sources or publications, but we can provide some references.
Source material for films, cinemas, and companies will hopefully exist for many of your databases. I have opted to leave this out of the schema. I do think that they are important, and should be included whenever possible, but when merging databases, I think it's best to deal with just the "data", and allow the sources, to be browsed, or exchanged through a separate process.

Company

The following was suggested as a way to indicate the relationship of a company to various other entities within the database. The first three connections exist within the current database structure. The last 3 are not – please discuss the importance of the different relationships, or if you would prefer to indicate them in a simple text field related to each field/

Company relationships:

"A job is expressed by a link between the company and a venue, a movie, another company, a person, an address, etc.. Some information has to be added as to the meaning of the link. My suggestions:

1. A link with a venue means exhibition by a cinema chain.
2. A link with a programme means exhibition by a travelling cinema.
3. A link with a person means employment or ownership.

4. A link with an address means ownership of the building located at the address.
5. A link with a movie means distribution (production being outside this domain).
6. A link with another company means a mother/daughter company."

Company Name

"Significant part of the project is on rise and fall of exhibition, distribution and production companies pre=1927."

"Yes. I am collecting data mainly on distribution companies. When I collect programme data, I also try to find out production companies. If an exhibition venue is affiliated with some larger exhibition concern I also capture this info."

"Yes, it allows to draw networks of companies and cinemas"

Alternate Name of Company

"Might be important for various reasons, one of them being that we combined the structural database with oral history (interviews), and people often used other names for specific venues."

"Yes, names change and without this, the reconstruction of cinema history is impossible."… "It is sometimes difficult to decide when a company changes its name, or when a new company is established."

"Company names have official and unofficial variations, abbreviations etc."

Start Date/ End Date   First/ Last recorded existence of company

"Yes. I use start dates to trace the development of renting companies of different types."

"Dates are needed for identification. Some company names return after many years, so dates are one way to prevent a misidentification."

"Significant part of the project is on rise and fall of exhibition, distribution and production companies."

"I do capture info such as bankruptcy trials, voluntary winding-up processes, and receiverships. Again, this will be useful to date-stamp the entities that I intend to map, to

make database queries more consistent, and to produce more accurate snapshots of relationships between companies."

Info   A long text field to allow other descriptive information about the company

"Yes, for instance to refer to the sources for the information you use! Might be useful if you want to understand 'strange' trends in your database."

"Yes, a lot of the info I find about companies is not easy to classify and it makes more sense to just store it there to be retrieved when needed, eg when writing more descriptive parts of essays."

"A short note may do in most cases. In my view, details about a company should function like an encyclopedia. You want to know enough to be able to identify a company, and link it to a job in moviegoing history."

Job + Date   The Company's relationship with the Theater
(supplier, production, financier, etc – need controlled vocabulary)

"We used it, but we didn't work enough with this kind of different data yet."

"Useful information for the business history part of the project."

"Yes. I am trying to figure out which distribution companies supplied which cinemas, so I try to capture any indication of this in a text field because at the moment the info is rather vague and inconsistent."

"Yes it enabled us to see the different layers of power in the exhibition sector- somes companies, not only owned a lot of theatres, it also programmed in theatres it did not own."

"I have used job titles to prepare a dataset for social network analysis."

- Have you ever used this info in your research, and if so how. If not, are you likely to? If so how?

OTHER NOTES ABOUT COMPANY
- Type of company, ie main object: producer, distributor, exhibitor, and possibly several categories within each area, eg. Distributor could be agent, second-hand dealer, renter, film service, state-rights/exclusive manager, etc.
   > **This type of information should be covered within the connection of the company to the cinema through the job.**
- People involved: directors, main shareholders
   > **A person can be connected to a company Through the Person – Job – Company relationship.**

<u>Person</u>

Name   First name, last name.
"Yes-I've tracked who managers, programmers, and other personnel were at some cinemas."

"Occasionally useful, but not essential."

"Yes. Individuals are useful indicators of relationships between other entities. For my current project I am keeping the names of cinema proprietors, cinema managers, traveling exhibitors, and managers of renting companies or hiring departments of larger companies. This allows me to track multiple ownership, formation of cinema chains, movement of personnel between cinemas, etc."

Alt. Name   to allow for alternate spelling of a name.
- Is this an essential bit of information for your research?   YES NO
- Have you ever used this info in your research, and if so how. If not, are you likely to? If so how?

Birthday/ Death date   YYYY-MM-DD, YYYY-MM, YYYY
                              Or any amount of data known.
These bits of information seem to be used anecdotally, and to add additional information, but not essential..

"Family histories and network histories are often very complex."

"To situate the person"

"I have not used it and am not likely to use it. It would be more useful for me to know when an individual entered the film trade, regardless of age."

"I have not used it, although I try to keep it if I find it. I do think it is useful because someone's death causes rearrangements in business organizations, and can be a marker of generational change."

Place of Birth   City, State, or Country
"I have gathered some of this information informally, but not much."

"I try to make a note of birthplace if I find it mentioned, but this is not essential to my research. Again, knowing where the individual started trading would be more important."

"No, not really relevant, unless to specifically situate the person"

Race/ Ethnicity/ Ethnic Origin/ Religion/ nationality)?
                              "Probably yes, although would that not impose dubious
                              classifications according to contested notions of ethnicity?"
Race does not seem to be "essential" for everyone, but if this term were broadened to include – ethnic origin/ religion/ nationality, would it them be more useful?

"In some sense (I'm sorry for this..) we did, where the network was controlled and run by a rich local businessman (at least if this falls under the category of racial distinctions…, but origin is important in this case.)"

"Definitely-an important part of my attempts to understand race in relation to filmgoing."

"I have not used it and I am not likely to use it, as it does not seem such an important factor in the context I am studying. I do keep notes on ethnic origin if known, but I do not actively try to establish it."

"This type of information is controversial in the public domain. I am not sure whether or not to include this in a public database. At the same time, I won't stop a researcher who wants to analyze Jewish ownership of cinemas. I wonder if this element is also part of standard metadata of persons."

Biography/ Notes   A long text field to allow other descriptive information about the person.

"Not essential, but occasionally useful in identifying key figures and family connections in travelling showpeople.

"Yes. I add notes from different sources to build a fragmentary chronological sketch. I try to see similarities and differences in people's life trajectories. This is not a formalized approach, it is mostly used as context."

"A short line may do in most cases. In my view, details about people should function like an encyclopedia. You want to know enough to identify a person, and link this person to a job in moviegoing history."

Job   The job title + Beginning & Ending Date (?)
It seems that it is important to situate the relationship of a person to a company or cinema – would this be a field where an editable controlled vocabulary might be handy? So that everyone can be using similar terms, or are they far too varied?

"Yes. For most cases, for example cinema managers, this relationship to a cinema is the only info about them that I retain; they exist mostly as attributes of the venue, and as connectors between several venues. It is also interesting to see when new job titles emerge, such as 'programme controller', 'circuit manager', 'dispatcher', etc."

"I have used job titles to select the management of cinemas, preparing a dataset for social network analysis.
A job consists of a link between an employee/person and the company or venue. The job title is a qualification of the link. Someone's profession can be mentioned in the Biography field."

Place of Death
"Place of Death is required for identification purposes."
- Is this an essential bit of information for your research?   YES NO
- Have you ever used this info in your research, and if so how. If not, are you likely to? If so how?

Gender
"Sex should be included to fit standard metadata."
- Is this an essential bit of information for your research?   YES NO
- Have you ever used this info in your research, and if so how. If not, are you likely to? If so how?

Address
Town or country of residence. (what level of detail might be important?)
"In some instances, country of origin will be used in our project - particularly to distinguish indigenous involvement in cinema business from inward investment from external investment."
- Is this an essential bit of information for your research?   YES NO
- Have you ever used this info in your research, and if so how. If not, are you likely to? If so how?

Ideological Groups (the ideological groups which this person may be associated with)
- Is this an essential bit of information for your research?   YES NO
- Have you ever used this info in your research, and if so how. If not, are you likely to? If so how?

OTHER NOTES ABOUT PERSON

"Previous situations: This would not have to be a field, but just the ability to see all the records that contain that person's name in chronological order."

> **This could be accomplished by querying the Person – Job – Company relationship or the Person – Job – Cinema relationship. I have noted to add date to both Job relationships..**

"Relationships to other members of the trade (eg brother of…, married to…)
Any thoughts on the importance of building this into the database, or could it just be included with in the Biography?

"Archival references – where it would be possible to find out more about this person
> **Source material for films, cinemas, and companies will hopefully exist for many of your databases. I have opted to leave this out of the schema. I do think that they are important, and should be included whenever possible, but when merging databases, I think it's best to deal with just the "data", and allow the sources, to be browsed, or exchanged through a separate process.**

Program
A collection of shows for preview in a specific Cinema

I have adjusted the DB Schema so that a programme connects directly to the cinema – a program would be a collection of films, or a film schedule within particular cinema. If you definition differs, please indicate.

"Regarding Programme definition:
A programme describes what could be seen, where and when. Basically, it lists a number of movies and live acts, in a certain order. The individual items of a programme can be seen as performances, live as well as recorded (in a silent movie you'll have both at the same time). Musical accompaniment and lecturing are part of a programme, usually a job of employees, but sometimes announced as a special contribution. The latter could be mentioned in a note."

Program Title
"In my experience, programmes do not have titles. They could be given a title corresponding to the 'feature' film, but this would only work for some historical periods."

"Special programmes are announced as such in the press, like gala premiere, Sunday morning event, late night show, Christmas show, and most importantly: matinees for children or teenagers. I have used these programme names/titles often to include or exclude records for further analysis, i.e. to prepare a clean dataset."

Program Alt. Title
"This refers to alternative titles, but also looking for the original title: Some newspapers mostly only referred to translated titles: not easy to find the original ones!"

"A single title (of the programme) should be enough, I think. The main title can have a subtitle, however."

Program Date Date of this specific Program
"We mainly worked on a weekly basis."

"Crucial for tracking understandings of distribution patterns, patterns of availability."

"Tracking distribution of programs."

"Yes. The dates of film shows have different uses in my research. I can think of a few now, but there might be more: a) they allow me to see when programming practices changed, eg from the 'variety' type to the 'feature' type of programme; b) by date-stamping the data and then plotting it to a map, I will (hopefully) be able to see how films traveled around the country; c) by comparing the programme to a release schedule I can see how far a cinema was from being a 'first-run' house; d) by comparing the programme to different renters' offers, I hope to be able to infer where a particular cinema was getting its films from."

"I have used dates of film shows for analyzing the supply, distribution and scheduling of movies in cinemas, cities and regions. In our database the week is the smallest measure of time. We always register a Friday to mark the beginning of the weekly programme; a

continuation into the next week requires an extra date for the next Friday. A programme change during the week calls for a new record while additional details about hour and day can be specified in a note. It allows me to select a specific period of time, and to count the number of Fridays a movie was screened in one or more cinemas. In doing so it is not necessary to specify a date of closing."

Program Category
This seems to be primarily useful if it's a mixed program.
Should we limit this to just indicating whether it's a mixed program or only film?

"I'd prefer not to indicate genre, since I'd like to let researchers make those determinations for themselves-though I think the indication of "mixed program" (meaning combining live and filmed programming or features/shorts/newsreels) is useful."

"Particularly interested in mixed programs "

"Film metadata may already include details about genre, so it seems a waist of time to specify film genre in a programme description."

Distribution Company The Company that distributes the Programme
This is the relationship of Company – Distribution – Program
By adding a Job Title & Date to this relationship, it allows you to define the relationship of a company to a program in any way you like.

"If we found information on this we included the data. But we mainly worked with film program overviews in newspapers & they mostly didn't refer to distribution companies or employees."

"Yes. This is one of the main goals in my project, but it is notoriously difficult to establish which companies were supplying films to a given theatre, especially in the early/transitional period when there were no block-booking contracts, and in an open-market system. So far I focus on following the trail of single films whose initial distribution companies I know. I want to explore the geographical patterns of distribution – which companies supplied which areas."

"Yes, to determine the programming strategy of the cinema/company, and the distribution networks"

"Distribution information belongs to the movie, not the programme. I would suggest to change this item to Travelling Cinema Company."

Distribution Employee A person that distributes the Programme
This is the relationship of Person – Distribution- Program, or how a person is connected to a company that can then be connected through the relationship of Company – Distribution – Program
By adding a Job Title & Date to this relationship, it allows you to define the relationship of a person to a program in any way you like.

"Will be important in some instances in establishing travelling showpeople"

"Yes, but I keep this info as an attribute of the distribution company."

"I don't understand what a distribution employee has to do with a specific programme in a local cinema. This person is working in a distribution company. I would rather see this job as a (steady) relationship between the cinema and the distributor, which can be traced by checking the movies on show in a cinema, but it does not show up in the description of a single programme. My suggestion is to change this element to refer to a temporary employee (not a performer) involved in this particular show, like a local assistant of a travelling showman."

Price
"Special programmes call for special entrance prices. Though it is not necessary to mention price in every programme, it is useful to record entrance prices at intervals."
- Is this an essential bit of information for your research?   YES NO
- Have you ever used this info in your research, and if so how. If not, are you likely to? If so how?

OTHER NOTES ABOUT PROGRAMME

"Project is interested in how early cinema was reviewed in local press. Whether or not a review has been identified would be useful information."
> **Source material for films, cinemas, and companies will hopefully exist for many of your databases. I have opted to leave this out of the schema. I do think that they are important, and should be included whenever possible, but when merging databases, I think it's best to deal with just the "data", and allow the sources, to be browsed, or exchanged through a separate process.**

"It might be useful to know whether it was promoted in any special way, for instance as an educational show, or as a Sunday special show."
> **Can this be covered within the source materials, or notes, or tags?**

"Order -The order of movies and live acts in a programme an important characteristic of early cinema. The database must allow to register the order number of each performance if known."
> **I've added order to the relationship of Film – Part of – Program, this way with the relationship can also indicate what order the film was show within a program.**

"Live Acts - Live acts by artists can be part of a cinema show. They are more difficult to identify and standardize than film titles. Our database has a field for a brief description of the artist(s) and the act without further identification, and its order number"
> **How important is the addition of Live Acts? Could this be added as a "Film", and have the attributes of a film, but perhaps with a notation of "Live Act" as the Genre??**

Film

Film Title

"Particularly interested in movement from programmes to features."

"Yes. Tracking the exhibition of a particular film in different places is an approach I am exploring to try and tease out distribution patterns."

Film Alternate/ Translated Title
- Is this an essential bit of information for your research?   YES NO
- Have you ever used this info in your research, and if so how. If not, are you likely to? If so how?

"Please allow for alternative and translated titles."

"perhaps alternate could also be "translated"

Film Release Date   YYYY-MM-DD, YYYY-MM, YYYY
        Or any amount of data known.

"Yes, to determine the programming strategy of the cinema/company"

"Yes. I need to know the release date of a film so I can observe how long it took to appear at a given cinema."

Film Production Year
"This might be important to understand programming strategies: how "old" are the movies played in a theatre?"
- Is this an essential bit of information for your research?   YES NO
- Have you ever used this info in your research, and if so how. If not, are you likely to? If so how?

Country of Production
"Key information when trying to identify indigenous production. Not only country but place"

"I have not used this info yet, but as different forms of film trade operated between different countries, it is important info for my project. Films did not come in the same way from the US to the UK, as from France to the UK; different people were involved, different transactions, etc."

"Yes, to determine the programming strategy of the cinema/company and the characteristics of a cinema based on origin of the films."

Directed by (I've left this as a plain text field & not connected to "people", as people would indicate people involved in film ditribution and cinema work – is this acceptable?)
"May be useful rather than essential, but would be included where it is significant."

"This would be only useful to add some anecdotal info, or to differentiate between films of the same title, but the film text is not really important in my project, and the notion of

director was only starting to take shape in the period I am studying, so I do not think I would use this field."

"To determine the film."

"Yes, I have used information about directors and stars to select a record set, though this information is not in our database. First, I prepare a list of movies by searching IMDB. Next, the list is used to find the relevant titles in our database."

Production Company (is this ok being a plain text field not associated with the other companies in the database – thus would carry no attributes, but allow for a place to enter the name of the production company?)
- Is this an essential bit of information for your research?   YES NO
- Have you ever used this info in your research, and if so how. If not, are you likely to? If so how?

Film Length

"Key data for tracking the emergence of the feature film."

"Yes. Long films used distribution channels that might differ from those established for short films. I think the adoption of the long film programme marks an important separation between cinemas in around 1914-5, so knowing if a cinema was showing long films, and if distributors were supplying them, is crucial historical info."

Film Genre

"Genre is crucial for understanding programming strategies, but incredibly tricky, as you probably understand. We used all kinds of external databases in order to include generic data into our programming database (e.g. generic descriptions from IMDb), but this proved to be unusable. What we finally did is to build a genre category system based on the labels given by the newspapers and the adds at the time when the movies came out! The genre labels often strongly differed from how we (or IMDb users) define movies today. This is an interesting topic in itself (on which I would like to work, once …)."

"It is not essential but if I have the info I use it, at least to establish broad categories such as topical, drama or comedy. This is important in relation to 'variety' programming."

"Our database only uses fiction, nonfiction, animation, and series."

Film Censorship
- Is this an essential bit of information for your research?   YES NO
- Have you ever used this info in your research, and if so how. If not, are you likely to? If so how?

"We also included <u>censorship data</u>"

Language
- Is this an essential bit of information for your research?   YES NO
- Have you ever used this info in your research, and if so how. If not, are you likely to? If so how?

"We also included <u>language</u> items"

"Dubbed into x language"

"Subtitiled?"

Link to IMDB record/ or other ext reference
- Is this an essential bit of information for your research?   YES NO
- Have you ever used this info in your research, and if so how. If not, are you likely to? If so how?

"External references, especially whether the film exists in some archive"

OTHER NOTES ABOUT FILM

Distribution model, ie whether it was an open-market feature or an exclusive (in the UK); a licensed or independent release (in the US)

Source material (novel, stage play, vaudeville act, original screenplay, etc)

> **Source material for films, cinemas, and companies will hopefully exist for many of your databases. I have opted to leave this out of the schema. I do think that they are important, and should be included whenever possible, but when merging databases, I think it's best to deal with just the "data", and allow the sources, to be browsed, or exchanged through a separate process.**

Screening hours/days/weeks: to reconstruct the path of a specific film over cinemas/cities

> **Screening hours, days, weeks should be part of the film's program at a specific cinema.**

~~Film Episode~~

It seems not many people use Film Episodes – as they relate to serials from the time of early cinema – so I have deleted this section. Please let me know if denoting a separate table for film episodes is important to your research.

Appendix 14: Tag Survey (Delphi Round 2)

1. Introduction.

These tags are a cummulation of all of the tags that had at least 1 vote from the last round, as well as any suggestions. Please go through and select all of the tags, which you feel would be helpful to be able to compare/ share across the various projects.

These terms would be used as descriptive "tags" for your primary sources, movie venues, companies, or anything else describe within your system. The idea is to come up with a common set of terms accross collections.

2. Tags

Please select any of the terms you would find relevant to your research.

1. Please select all that would apply to your research/ interests.
Again these "tags" would ideally be added to primary sources, movie theaters, people, or any other item which is being describe by the system

- "talkies"
- actualities
- admission price
- advertisement
- age
- air conditioning
- airdome
- Amateur Show
- American indians
- amusement park
- architect
- Audience
- balcony
- balcony, racially segregated
- band or orchestra
- black theater
- Building Cost
- carnival
- censorship
- censorship data films shown
- change in owner
- charity
- children
- church
- cinema chains

- electrical generator
- electricity from mains
- ethnicity
- event
- fair
- family
- fashion show
- feature film
- Female Exhibitors
- fight films
- film entrance
- film societies
- film title
- fire
- foreign movies
- gas from mains
- gender
- geographical expansion
- high street / main street / downtown
- hotel
- hours of operation
- ideological
- illustrated song
- imports
- influenza
- intermittent exhibition

- morality
- motion picture party
- music
- music-hall
- name change
- neighborhood
- Newspaper Clipping
- newsreels
- non-theatrical venue
- notable
- open-air screening
- opening date (broader than opening notice?)
- opening notice
- opera house
- Oral History
- organ
- origin of films shown
- penny arcade
- Photograph
- pianist
- Postcard
- press reception of cinema
- program change
- projection problems
- projectionist

- circuit or chain
- circus
- closing
- color films
- company
- comparative studies of popularity and preferences
- concessions
- contest
- crime
- distribution
- Downtown
- dress code
- drugstore
- economic integration

(eg only one day a week, only special films)
- Labor Union
- ladies and children
- lecturer
- licensing
- local films
- location change
- manager
- Map
- marketing
- men only screenings
- metropolitan
- midnight shows
- minstrel shows

- public hall
- publicity stunts
- Railroad
- rear projection
- regional selections of movies
- Rural
- seating policy
- silent film
- social class
- specific films
- suburban
- Sunday opening
- Theater Chain
- WWI

3. Comments

1. Please include any comments about this part of the process.

Appendix 15: Final Data Dictionary

<u>Cinema</u>
(* indicates a "core field" – which is almost universally desired)

*Cinema name
The official name by which the Cinema is called

Screens + Date
The number of screens a cinema has at a particular time.

Cinema Seating Capacity + Date
Amount of Seating available
Actual number if available or:
0 – 150, 150 – 300, 300-599, 600-1500, >1500

Racial Policy
The racial policy of a particular cinema. This would ideally be a controlled vocabulary within each project.

*Address
The location of the cinema. While GIS coordinates are preferable, this is not always possible. When possible, please include street address, or at the very least, city, county, state, country – etc.
Location will be used not only to plot the cinemas on a map, but also for disambiguation where there are multiple cinemas of the same name.

*Opening/ Closing
The date a cinema opened and closed at a particular location. Where there is not exact data, and estimate is preferred. (e.g. ca 1970)

Ticket price range + Date
The ticket price range + date where noted.

Building type
If building's primary purpose is anything other than a cinema, note here what the primary function is.

Political Orientation/ Ideological Association
The official name of a political or ideological group that the cinema is associated with.

Technology + Date
The name of a "new" technology (sound, air conditioning, 3D, etc) + the date it was first mentioned/ installed.

<u>Relationships:</u>
Cinema – *Relationship (+ date)* – Cinema
      The Company's relationship with the Theater
      (supplier, production, financier, etc – need controlled vocabulary)
Person – *Relationship (+ date)* – Cinema

A person who works for a specific Cinema

Program – *Relationship (+ date)* – Cinema

A program that was shown at a specific Cinema

<u>Company</u>
(* indicates a "core field" – which is almost universally desired)

*Company Name
The proper name of the company.

Alternate Name of Company
An alternate name or abbreviation. Useful for common names or un-official variations

*Start Date/ End Date
First / Last recorded existence of company

Company Address
The address of the company – to as much specificity as possible. (this is not primarily to be used to map locations, but to identify and situate the company

*Info
A text field to allow other descriptive information about the company

Political Orientation/ Ideological Association.
The official name of a political or ideological group that the cinema is associated with.

<u>Relationships:</u>
Cinema – *Relationship (+ date)* – Company
  The Company's relationship with the Theater
  (supplier, production, financier, etc – need controlled vocabulary)
Person – *Relationship (+ date)* – Company
  A person who works for a company
Program – *Relationship (+ date)* – Company
  Distribution companies or other
Company – *Relationship (+ date)* – Company
  A company is related to other company

<div align="center">

<u>Person</u>
(* indicates a "core field" – which is almost universally desired)

</div>

Name
First name, {Middle name}, last name.

Alt. Name
to allow for alternate spelling of a name.

Birthday/ Death date
YYYY-MM-DD, YYYY-MM, YYYY
Or any amount of data known.
        Mainly used to situate the person.

Place of Birth
City, State, or Country

Race/ Ethnicity/ Ethnic Origin/ Religion/ nationality?
This is possibly controversial to publish this field to a public database, but it seems important for research. So if able to gather the information, it would be useful, even if you don't display it.

Biography/ Notes
A text field to allow other descriptive information about the person.

Gender
M/F

Address
Street Address if possible; Town or country of residence

Ideological Groups
(the ideological groups which this person may be associated with)
Ideological associations, Labor unions, etc.

<u>Relationships:</u>
Cinema – *Relationship (+ date)* – Person
        A person who works for a specific Cinema
Company – *Relationship (+ date)* – Person
        A person who works for a company
Program – *Relationship (+ date)* – Person
        Traveling Showpeople or other

<u>Program</u>
(* indicates a "core field" – which is almost universally desired)

A collection of shows for preview in a specific Cinema

a program is a collection of films or other acts, or a film schedule within particular cinema. If you definition differs, please indicate.

Program Title
It is understood that not all programs have titles, but this would be a descriptor for a special program such as but not limited to: Sunday morning event, late night show, Christmas show, and most importantly: matinees for children or teenagers

*Program Date
Date of the program to as much specificity as possible

Mixed Program?
y/n

Price +date
The price of a specific program. Perhaps recorded at intervals?

<u>Relationships:</u>
Company – *Relationship (+ date)* – Program
        Distribution companies or other
Person – *Relationship (+ date)* – Program
        Traveling Showpeople or other
Cinema – *Relationship* – Program
        A program that was shown at a specific Cinema

<u>Film</u>
(* indicates a "core field" – which is almost universally desired)

*Film Title
The official title of the film

Film Alternate/ Translated Title
An Alternate or translated title of the film

Film Release Date
YYYY-MM-DD, YYYY-MM, YYYY
Or any amount of data known.

*Film Production Year
The year the film was produced

*Country of Production
Country the film was produced in.

Directed by
Text field to put director's name

Production Company
Text field to put Production Company's name

Film Length
Length in minutes of the film
Eg. 1:45 (One hour forty five minutes long)

*Film Genre
The genre of the film – should be a controlled vocabulary within each individual project.

Film Censorship
Film rating, or censorship information

Language
Language of the film
Dubbed, subtitled, etc.

Link to IMDB record/ or other ext reference
IMDB url or other external reference.