Genre Containers: Building a Theoretical Framework for Studying Formats in Information Behavior

Brittany Brannon, https://orcid.org/0000-0001-7090-3342, OCLC Research, 6565 Kilgour Place, Dublin, OH 43017, brannonb@oclc.org, Corresponding author.
Amy G. Buhler, https://orcid.org/0000-0002-0511-8273, Marston Science Library, University of Florida, PO Box 117011, Gainesville, FL 32611-7011, abuhler@ufl.edu
Tara Tobin Cataldo, https://orcid.org/0000-0003-4569-6374, Marston Science Library, University of Florida, PO Box 117011, Gainesville, FL 32611-7011, ttobin@ufl.edu
Ixchel M. Faniel, https://orcid.org/0000-0001-7302-5936, OCLC Research, 6565 Kilgour Place, Dublin, OH 43017, fanieli@oclc.org
Lynn Silipigni Connaway, https://orcid.org/0000-0001-9783-675X, OCLC Research, 6565 Kilgour Place, Dublin, OH 43017, connawal@oclc.org
Joyce Kasman Valenza, School of Communication & Information, Rutgers University, 1277 Cox Road, Rydal, PA 19046, joyce.valenza@rutgers.edu
Christopher Cyr, https://orcid.org/0000-0001-5246-6900, OCLC Research, 6565 Kilgour Place, Dublin, OH 43017, ccyr85@gmail.com

Note: This is a preprint of an article published in Journal of the Association for Information Science and Technology, copyright 2021. This preprint has been updated to reflect changes in the final version.

This preprint is deposited on the OCLC Research website under the Creative Commons Attribution Non-commercial International License 4.0 (CC BY-NC 4.0).

Suggested citation:
Abstract
Prior studies have shown high-level differences in people’s perception and use of various information formats. However, the lack of a coherent and theoretically informed framework of elements of format has inhibited a nuanced understanding of the role that formats play in information behavior. This paper draws on theories from the field of rhetoric and composition to ground the study of information format in a social constructivist perspective that foregrounds action in context. Specifically, rhetorical genre theory is discussed in detail and the limitations of previous information behavior studies using rhetorical genre theory are explored. One of the main problems of earlier studies is in confusing genres and their containers. This paper introduces and defines the concept of containers as typified ways of collecting and presenting texts of certain genres for publication. Building on rhetorical genre theory, the paper offers a theoretical exploration of the role that containers play in the formal and/or public sharing of information within discourse communities. An illustrative example of the concepts applied to data from an IMLS-funded study is provided.
Introduction
The emergence of the internet and the shift to digital communication disrupted the ways that information was created, published, and circulated in print. One effect of this was that the physical structure of information objects disappeared, a phenomenon called container collapse (Connaway, 2018; Valenza, 2016). Educators and information professionals now question whether information format plays an important role in information behavior. Williams & Rowlands captured this belief in 2007 when they reviewed the literature to try to determine whether or not young researchers are “format agnostic” (Williams & Rowlands, 2007, p. 20). Williams and Rowlands found no empirical evidence to support this belief, but it is an assumption that has continued to inform both research and practice. A few researchers directly answered their call for research into this “hugely important issue” (Williams & Rowlands, 2007, p. 20), with studies suggesting that students are not format agnostic (Cataldo & Buhler, 2012; Francke & Sundin, 2009). However, this area remains understudied and more research is needed to understand the role that format plays in the creation, seeking, evaluation, and use of information.

One challenge in studying the role of format in information behavior is the lack of a clear, theoretically informed approach for describing information formats. Gorichanaz (2017a) found that there was significant confusion among library and information science (LIS) professionals over what the terms genre, format, and medium meant and how to deploy them. Most studies of format in information behavior take format terms and concepts for granted, discussing them without defining or explaining them. Consequently, this paper draws on rhetorical genre theory to provide an established and robust theoretical foundation for discussing elements of format that are grounded in social practices of communication.

Rhetorical genre theory is focused on genres of nonfictional prose, and defines genres by the social actions that they allow. From this perspective, genres and containers are, first and foremost, sets of socialized and internalized communicative practices for acting in the world by creating, sharing, interpreting, and using information. These practices result in more or less similar types of information resources, and it is the recognizability of those types that help to mediate between creation and use and between people and social structures. While these practices result in information objects that embody the process of their creation, it is counterproductive to view the objects themselves as having primacy. The recognizable formal elements are significant only in the context of the social and cultural communities and systems of meaning in which they participate.

This paper begins the process of creating a precise, descriptive framework of elements of format by drawing on rhetorical genre theory to define containers. Containers, as typified ways of collecting and presenting texts of certain genres for publication, are an important element of format to consider when studying information behavior. Container and genre, along with mode, file type, and medium, are presented as a framework that offers a nuanced view of format in information behavior.

Literature Review
Format in Information Seeking
One of the primary areas of focus for studying format in information-seeking behavior has been users’ preferences for print vs. electronic materials. The dawn of the internet era significantly changed the way people engage with information (Burke, 2010) and the types of materials that libraries needed to provide to their users (Zauha & Ragains, 2011). To adapt, library and information science (LIS)
professionals had to better understand when, how, and why their users were using resources in each medium. Some studies look at the preferences of specific user groups for physical and digital materials (Chen, 2019; Clark et al., 2018; Dukić & Strišković, 2015; Superio et al., 2019). Other studies look at the use of specific material types, such as reference materials (Chan & West, 2015) or primary sources (Press & Meiman, 2020), in physical and digital forms.

One major thread of this research focuses on e-book usage. Many of these studies examine specific subgroups to determine when and how they use e-books (Dahl, 2013; Hartel & Cheek, 2011; Pinto et al., 2014). Recently, Tracy (2018) called for a more nuanced approach to the category of e-books, arguing that what most scholars treat as a “unified experience” is actually a “set of formats” (p. 20). He provides the term microformats to capture the variety of technologies, including browser-based interfaces, pdfs, and printouts, through which users might experience e-books (Tracy, 2020). He finds that users engage with e-books in different microformats depending on the information-seeking task that they are trying to accomplish. One major limitation of the study is that the term microformats is not defined and is instead explained by listing the specific microformats under discussion. The lack of definition makes it difficult to determine what would or would not count as a microformat and limits the ability to apply the concept of microformats to other types of information resources. Both the nuance of Tracy’s findings and the limitation of the terminology support this paper’s call for a more structured and theoretically informed approach to describing formats.

Another area in which the issue of format has drawn attention, albeit superficial, is in the discussion around the implementation of federated search and discovery layers. As early as 1985, librarians were discussing format as an element of search strategies for helping users navigate information systems (Quint, 1985). As single-search interfaces have become more popular, library and information professionals have debated the effect on users of receiving results in a variety of formats. Haggerty and Scott (2019) call it a “calculated risk” of these systems to reduce complexity at the point of searching but increase the complexity of evaluating search results, including the need to distinguish among a variety of format types. Holmes et al. (2008) found that school-age children struggled to differentiate among the formats returned by their search. Greer and McCann (2018) found that undergraduate students do not recognize the differences between online formats, and struggle to cite sources for their papers. There is also evidence to suggest that students are not using the format filters within the discovery system to navigate these complexities (Resau, 2019).

For some, these new interfaces represent both a need and an opportunity to change information literacy instruction (Fawley & Krysak, 2012). Some have created information literacy instruction sessions that directly address this issue, either in the discovery layer only (Perry, 2014) or by using the discovery layer in conjunction with specialized databases (Scott, 2016). Seeber (2015) contends that teaching “format as a process” necessitates moving away from a focus on discovery systems themselves to focus on the processes by which the information resources within them are created.

More recently, the nature of preprints as a format, and their relationship to the scholarly journal article, has become a topic of interest (Kim et al., 2020; Klein et al., 2019; Pagliaro, 2021). Despite this interest, there is still relatively little research examining how preprints are used (Chiarelli et al., 2019). Equally understudied is the way that information seekers view preprints or whether they recognize the differences between preprints and published journal articles (Cataldo et al., 2021). The COVID-19 pandemic heightened attention to and debate over the role that preprints play in scholarly
This literature confirms that today’s users are not format agnostic. The studies demonstrate broad differences in the use of format types, but it is difficult to draw nuanced or consistent conclusions. It is also worth noting that in most of these studies, format is approached in a common-sense way. That is, authors do not explain why they chose a format, how they define that format, or how they know that the examples they give are examples of that format. This is one of the challenges in building a more systematic body of knowledge around the role of various format types in information behavior. Additionally, most of these studies are observational studies of use or reported preference. The extent to which users are able to distinguish between formats or understand the significance of formats is still largely understudied.

Rhetorical Genre Theory in Information Behavior

Jack Andersen has made numerous arguments for the use of rhetorical genre theory as a framework for understanding information seeking and evaluation. He suggests that “a particular genre defines an information need” because “a genre, and the form of knowledge it materializes, is defined and constrained by the activity it is used to accomplish” (Andersen, 2008, p. 33). In literate societies, the ways in which information is created and transmitted are inseparable from society’s structures, values, and practices (Andersen, 2006). Genres, as typified discursive actions, are the nexus of these broader social structures and individual activity.

Since Andersen, rhetorical genre theory has gained traction in LIS. Scholars have looked at genre-labelling of webpages (Santini, 2008), the use and implications of rhetorical genre theory in conceptualizing scholarly communication (Gullbekk, 2016; Nuria, 2011), disciplinary communities and specific genres within those communities (Gorichanaz, 2017b; Huvila, 2019; Nahotko, 2016; Schreiber, 2014), and applying rhetorical genre theory within information literacy instruction (Burkholder, 2010; Jankowski, 2018; Leebaw, 2018; Mills, 2014; Simmons, 2005). Researchers have combined rhetorical genre theory with existing theories and frameworks for information behavior (Gorichanaz, 2017b; Gullbekk, 2016; Huvila, 2019; Schreiber, 2014), but none have used rhetorical genre theory to form a coherent theoretical framework for talking about format in information behavior.

Limitations of Previous Studies

One area where format identification and understanding has been studied is users’ ability to recognize genres. These studies have found that users rely on genre to guide their information use (Francke & Sundin, 2009), but struggle to apply genre labels and often disagree on the labeling of the same information resource (Leeder, 2016; Rosso, 2005; Santini, 2008). What none of these studies question is whether the difficulty of users in applying genre terminology is the result of the terminology itself. Despite citing rhetorical definitions of genre, these studies do not explain how or why the genre terms are chosen and whether the “genres” presented to study participants are valid. In part, this is because format has been undertheorized in library and information science. A look at how three LIS studies deployed the concept of genre will help to illustrate the challenges. First, however, a brief reflection on the use of the term format will provide background for the conversation.

Williams and Rowlands’ (2007) review sought to “elicit information on the information behavior of young people.” Among the claims they addressed with the literature was that the “Google Generation
are format agnostic and have little interest in the containers (reports, book chapters, encyclopedia entries) that provide the context and wrapping for information ‘nuggets’.” However, it is unclear what they mean by format and by container, or whether they intend these two terms to be synonymous. They do not define either term or provide citations that might suggest a definition. When discussing the potential ramifications of the format agnostic stereotype, they focus primarily on presentational differences between online and offline formats. They note, for example, that preferences for online versus offline reading vary depending upon the type of the book and the way it is being used. These types of differences are issues of the medium, or the channels of communication through which information is delivered. However, the examples that Williams and Rowlands give in the quoted statement indicate genres. This general confusion among format, genre, and medium is widespread among LIS professionals (Gorichanaz, 2017a).

Both within and outside of LIS, format as a concept has two primary uses, as technological units or as a catch-all term. Some use it to indicate technological units that package information objects within a given medium or container. This is the sense of the word that Sterne (2012) employs when exploring the history, development, and use of .mp3 files, similar to how Tracy (2020) uses the term microformats in his discussion of e-books. Genette (1997) briefly discusses the technological development of the book and ends by distinguishing between softback and hardcover books as the most common formats. Even within these uses, it is possible to question whether the term is being applied consistently. Sterne explicitly restricts the term to things that prescribe how an analog mechanism or digital technology can store and deliver content, seemingly precluding the use of the term for books.

Most commonly, however, format is used as a catch-all term to denote any formal element of an information resource that varies according to recognizable types (e.g., ACRL, 2016; Clark et al., 2018; Francke & Sundin, 2009; Greer & McCann, 2018; Quint, 1985). In this paper, we will use format in the more common general sense, with the understanding that in doing so it is necessary to articulate the more specific elements of format that might fall under that umbrella. We use the term file types, discussed below, to denote format in Sterne’s sense.

Case Study 1 – Discourse Communities
Leebaw (2018) examines the historic bias against business information at liberal arts institutions among librarians teaching information literacy. Finding a lack of use, she makes a case that “business information is a powerful genre for teaching core concepts” (Leebaw, 2018, p. 301). However, despite calling business information a genre, she defines business information by the genres that comprise it: “primary sources produced by, for, and about businesses, such as financial results and filings, market research, trade publications, and marketing communications, such as websites, advertising, and public relations” (Leebaw, 2018, p. 301). In this case, the discourse community is being mistaken for the genre. Calling business information a genre creates the illusion that such information is monolithic. To truly understand the rhetorical practices of the business community, which she lists as a goal of teaching this type of information, it is necessary to see and understand the different roles that various genres play in enacting and shaping the practices within that discourse community.

Case Study 2 - Abstraction
Francke and Sundin (2009) conducted a rich ethnographic study of high school students’ work on group assignments. They conclude that genre is an important indicator by which students assess the credibility of resources. The categories that they list as genres are blogs, discussion forums, student papers, and
While student papers may be a genre, the other three examples are not. A student paper is a single, complete text. By contrast, a blog is a collection of blog posts, a discussion forum is a collection of posts in conversation, and Wikipedia is a collection of encyclopedia-style entries. Their analysis blurs the distinction between abstractions and collections. Genres are abstractions. They are conceptual models for texts based on people’s lived experiences and interactions with texts of that type. The other three formats discussed are ways of publishing collections of texts, or as will be discussed shortly, containers. Francke and Sundin’s analysis bears this out. They point out differences between students’ perceptions of certain container types and the specific information resources found within them. For example, some of the students held a belief that discussion forums were not very credible, but then found some individual posts to be credible based on their specific attributes. In this instance, the credibility of the container and the credibility of the genre may be at odds. This suggests that rather than students using genre to evaluate credibility, students use a combination of cues about the container type, the genre, and the particular resource they are evaluating to draw conclusions.

Case Study 3 – Unit of Analysis

Leeder (2016) took a quantitative approach when studying students’ genre knowledge, which complements qualitative ethnographic studies like the one conducted by Francke and Sundin (2009). He presented students with screenshots and asked them to apply one of the provided genre labels to each webpage. Unlike Francke and Sundin, who argued that students identify genres and use that identification during information evaluation, Leeder found that students struggled to identify genres. The labels that students could choose from were blog, book, book review, conference proceedings, database, encyclopedia, magazine, newspaper, research report, scholarly journal, and trade journal. These are all important types of information sources, but they are not all genres. Some are individual texts, whereas others are collections of texts. There is even, with the inclusion of the database label, an information system. The challenge here is in defining the unit of analysis. How should students apply a single label when multiple labels apply equally well depending on which unit of analysis they focus on?

The students’ least correctly identified information resource is illustrative here. It is an example of the book review genre, published in a scholarly journal, and located in the ProQuest database, making all three labels apply to it equally well. One student explicitly commented on this problem of nested categories: “Many of the screenshots could have been either databases or scholarly journals. The website was a database but there were often one or more scholarly journals one could access from that screen” (Leeder, 2016, p. 130). Leeder explains this as an example of the hybridization of genres online. While genres do hybridize, both online and offline, this is not the case here. In the print equivalent of this example, the book review would be housed in a physical journal issue, which would itself be housed on a shelf in the library collection. In this context, it is clear that the library collection is the information system which stores the physical journal, and the physical journal is the publication which contains the book review. These are not hybridizing genres, but three distinct elements of format that each individually plays a role in how a person might find, access, and use this information resource. In the digital world, the physical indications of these nested formats are flattened. That does not mean that when a book review is published online those three formats become hybridizing genres. What it does mean is that understanding and identifying these formats digitally poses additional challenges because the physical materiality of texts and information systems can no longer serve as a guide.

Leeder notes “Identifying the different formats in which information is published, i.e., books, magazines, journals, and newspapers, is also one of the basic components of traditional information literacy” (2016,
p. 126). He is absolutely right. However, it is important for information professionals to be clear about different elements of format and how they interact with each other to structure information and guide information behavior. We do not gain any explanatory power or analytical clarity by choosing to call all of these formats genres. Instead, we lose the ability to examine the ways in which information systems, publication types, and genres interact to shape students’ understanding and evaluation of texts.

**Project Background**

The IMLS-funded study *Researching Students' Information Choices: Determining Identity and Judging Credibility in Digital Spaces* (RSIC) examines the role that format plays in students’ evaluation of science, technology, engineering, and mathematics (STEM) information resources for a school research project (Cataldo et al., 2020). Simulated search results pages and a simulated research task provided data about the point-of-selection behavior (Buhler et al., 2019) of 175 students from fourth grade through graduate school, focusing on post-secondary students in STEM disciplines. The RSIC study focused on how students determine the credibility of relevant online content and the role that format plays in their decisions about helpfulness and credibility. Due to the general lack of clarity around elements of format in the information behavior literature, and the specific lack of a definition for containers, the project team drew on rhetorical genre theory. Rhetorical genre theory was a good fit because it focuses on genres of nonfictional prose, which are typically what students are relying on when conducting research in STEM. It is also a powerful theoretical approach due to its philosophical perspective on the nature of knowledge and communication.

Rhetorical genre theory is grounded in a rhetorical approach to communication, which views information not as an end in itself but instead as a socially situated tool that people use to achieve their ends. In other words, it foregrounds the fact that people use communication to *act*, both personally and socially. Day (2000) has argued that the historically document-centered focus of LIS, which resulted in the need to define, categorize, and control the objects that result from or facilitate human communication, has resulted in a hegemonic and normative view of communication as nothing more than transmission. This is the result of the unquestioning use of the conduit metaphor (Reddy, 1993) as the foundation for information studies: “The formal vehicle for perpetuating this, however, lies in characterizing language as a transmission and communication medium rather than as an agency for social, cultural, and political change” (Day, 2000, p. 811). This assumption creates challenges for understanding the intersections of format and information behavior, if formats are nothing more than neutral vehicles for transporting information content. The rhetorical approach to communication, in contrast, views communication first and foremost as created by and constitutive of social, cultural, and political engagement. In this view, formats emerge from and help to shape social and cultural practices, never neutral but always participating in these systems of value and activity. This view is also built on a social constructivist perspective which views communication, and therefore information, as existing in the individual construction of meaning based on shared social understandings. This is consistent with more recent LIS approaches to information behavior and information literacy, particularly with the adoption of the ACRL Framework for Information Literacy for Higher Education (ACRL, 2016).

Rhetorical genre theory was combined with elements of multimodality theory to define a framework of four elements of format for the study: container, genre, mode, and file type. Because container was of primary interest to the project, it was used in the development of one of the simulation tasks as well as in the qualitative codebook for the study. Genre, mode, and file type were used in the qualitative
codebook. These four elements of format combined allow the project team to parse the differences in students’ evaluation of information resources based on how they are constructed, structured, and shared. Medium is an important fifth element of format, but because the project focused only on information resources found and delivered through the Internet, variations in medium cannot be studied using project data (see Table 1).

**Table 1**

*Elements of Format*

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Container</td>
<td>typified ways of collecting and presenting texts of certain genres for publication</td>
<td>Book, journal</td>
</tr>
<tr>
<td>Genre</td>
<td>“typified rhetorical actions based in recurrent situations” (Miller, 1984, p. 159)</td>
<td>Article, chapter, editorial</td>
</tr>
<tr>
<td>Mode</td>
<td>“set[s] of socially and culturally shaped resources for making meaning” (Mavers &amp; Gibson, 2012)</td>
<td>Images, audio</td>
</tr>
<tr>
<td>File type</td>
<td>file format, or standard for creating and storing information in digital files</td>
<td>PDF, HTML, EPUB3</td>
</tr>
<tr>
<td>Medium</td>
<td>“the material form that carries the sign” (Burn, 2012); technological channel for communication</td>
<td>Internet, print, television</td>
</tr>
</tbody>
</table>

**Defining and Disambiguating Containers**

Based on the rhetorical definition of genre, this paper will provide a definition for containers, a theoretical examination of their role in the creation and sharing of information, and an example of the concepts in action. Genre, mode, and file type were all existing concepts, and the project team was able to use established definitions for them. One primary goal in developing this framework of elements of format was the disambiguation of these concepts. The rhetorical understanding of genres will be examined in depth to provide a foundation for both the container definition and other studies of format in LIS. After defining containers, this paper will also examine the ways in which containers are different from genres and provide a brief definition of modes and file types to distinguish those concepts as well.

**Genre Defined**

Within rhetorical genre theory, genres are defined as “typified rhetorical actions based in recurrent situations” (Miller, 1984, p. 159). This definition is frequently used and cited, within rhetorical genre studies and LIS studies using this theory, but it bears more thorough consideration because it is easy to misapply. It does not allow for any action to qualify as a genre, so it is worth outlining what type of situation and action are being described. It is also worth examining what is meant by typification and recurrence, and what effect that has on an understanding of genre.

The type of recurrent situation that Miller is discussing is a rhetorical situation, or a situation that can be addressed through the use of communication or persuasion. Not all situations will qualify. Running out of groceries is a situation best solved by purchasing more groceries, not by trying to persuade the empty cupboard to refill. Trying to get a new job, on the other hand, is an example of a situation that requires communicative solutions.
Similarly, the typified rhetorical action of Miller’s definition denotes a very specific type of action: a single, concrete communicative act. Let us continue with the example of trying to get a new job. Applicants write a resume to establish themselves as experienced candidates and write a cover letter to establish themselves as a good fit for the open position. Each of these rhetorical actions is part of the broader system of activity that is applying for a job. Applying for a job is not a genre itself because it comprises a variety of actions, some communicative and some not, some typified and some not, that involve a variety of actors, interactions, and systems. Resumes and cover letters, however, both represent single communicative acts within that system of activity.

Importantly, both resumes and cover letters are also typified communicative acts. Typification depends on recurrence. For rhetorical situations to recur, people must be able to recognize that a situation shares meaningful features with situations that they and others have faced in the past (Miller, 1984). As rhetorical situations recur, people repeat communicative strategies that they and others have used to successfully address similar situations. Over time, as successful strategies are repeated, they become typified into recognizable communicative acts that are not only successful but expected ways to address a recurrent situation. This typification means that genres are abstractions, or shared mental models of how to perform a type of communicative action. Mental models of genres are made up of implicit and explicit norms that guide how these communicative acts are created, shared, recognized, and received.

Having established the four parts of Miller’s definition of genre, let us now examine how genres achieve action and the role they play in structuring activity within communities. Figure 1 provides a model of genre action, or rhetorical action that takes the form of a genre, based in part on explanations provided by Devitt (2009) and Miller (1984). Genre action can be conceptualized most basically as a particular fusion of substance and form that creates an action that is recognized by others. Despite being represented separately, substance and form are inextricably intertwined (Devitt, 2009; Miller, 1984; Reddy, 1993). Miller defines substance “as the semantic value of discourse [which] constitutes the aspects of common experience that are being symbolized” and form “as the ways in which substance is symbolized” (p. 159). Devitt expands on this notion to include within form “all material embodiments of genre, linguistic and textual elements that might vary from one genre to another. Most obviously, then, form includes words, sentences, organizational structure, format, layout, and other visual elements” (Devitt, 2009, p. 33). Each of these various elements constitutes different ways in which meaning can be symbolized.

**Figure 1**

*Model of Genre Action*
This semiotic approach to form and substance acknowledges from the outset that despite the possibility of theoretically separating substance and form, in reality what is symbolized is inextricable from how it is symbolized. In other words, it is impossible to say that a particular form contains some type of meaning (Reddy, 1993). Rather, as in Reddy’s (1993) toolmakers paradigm, symbols are created based on one person’s experiences with the shared social meanings of symbols and then must be interpreted by another person based on their experiences with the shared social meanings of symbols. It is only through the interpretation of these symbols that form can be said to have substance, that is, that meaning can be made from otherwise arbitrary material signs. To change how something is symbolized, therefore, is to change what is symbolized. In this way, social and cultural contexts must be shared, at least to some extent, in order for symbols to be meaningful.

Another way that this can be viewed is through Skare’s (2009) argument that there are three ways to understand a document: as a physical phenomenon, as a mental phenomenon, and as a social phenomenon. While Skare assigns each of these phenomena to different disciplinary interests (documentation, information, and communication, respectively), they can also be understood within rhetorical genre theory where form is a physical phenomenon, substance is a mental phenomenon, and action is a social phenomenon. Skare starts from a document view, which assigns the document primacy as an object, independent of its social context. In contrast, because rhetorical genre theory starts from the fundamental view that all human expression is social, it would be misleading to suggest that form is only a physical phenomenon or that substance is only a mental phenomenon. Instead, both form and substance are fundamentally shaped by the social knowledge that makes possible the use of symbols to create meaning.

Together, substance and form make up the material expression, representing the concrete act of communication. This material expression is then interpreted by others as a type of social action, with the combination of the concrete act of communication and the interpreted social act making up the
rhetorical action described by genre. Communication in the social constructivist paradigm is interstitial, created between the intentions of the person using symbols and the understanding of the person interpreting them (Freadman, 1994). As such, a particular communicative action is not fully realized as a type of genre action unless it is recognized as such through another’s selection of and response to the pertinent genre features (Brannon, 2014; Freadman, 2002). A full understanding of genre action, therefore, requires attention to the situation that prompts the action, the intentions of the actor, and the contexts in which such action takes place (Devitt, 2009).

The individual’s intended response mediates between the rhetorical situation and the genre action. It’s important to note that intended responses do not necessarily equal genres. Rather, intended responses can be better thought of as the individual’s goals in communicating, or their rhetorical purposes. The conventions of genres can be thought of as raw materials and templates for achieving certain goals. These genre conventions can be mixed and matched, embraced or abandoned, as individuals utilize these social resources to accomplish their goals. Each instantiation of a genre will show variation in substance and form to a greater or lesser degree, depending on the fit between the individual’s rhetorical purpose and the genre(s) being enacted.

The combinations of substance and form that make up genre actions are only recognizable as actions because they occur within individual, social, and cultural contexts in which they are meaningful. Discourse communities provide one way of understanding how individual understandings are shaped, or not, by shared social and cultural contexts. Within discourse communities—groups of people who share aims, values, and communicative practices—communicative situations recur. Genres emerge when people within a discourse community find effective discursive methods of responding to recurrent situations. These discursive methods allow people to act in a way that makes their actions recognizable to other members of the community, and the repetition of the effective methods encourages other members of the community to adopt them. Eventually, they become naturalized into the structure of interaction in the community, an almost instinctive set of intentions and behaviors that members of the group can share. Genres thus act as a kind of mutual translator, identifying ways for individuals to achieve their goals through rhetorical action and guiding other people as they interpret and respond to those actions.

Devitt (2009) has argued that form is one of the primary mechanisms of this identification. However, she points out that form has been undertheorized in rhetorical genre theory, eclipsed by the emphasis on social action that was itself a reaction to formalist understandings of genre. In earlier works, Devitt pointed out the dangers of relying on form alone to define genre (Devitt, 2004), a caution she repeats by calling on scholars to acknowledge the materiality of texts without slipping back into formalism (Devitt, 2009). However, the elision of form in genre study has led to a sparsity of studies on the formal elements of texts that help people understand and define genres, and a terminological gap that can lead to confusion over what counts as a genre.

In the same volume housing Devitt’s call to reconsider form in rhetorical genre theory, one of the field’s most prominent scholars and a coauthor with expertise in digital rhetoric admitted that the focus on action had misled one of their previous analyses. Miller and Shepherd (2009) acknowledged that they were too quick to label the emergent technology of blogs as a genre. They point out that “with a rapidity equal to that of their initial adoption, blogs became not a single discursive phenomenon but a multiplicity” (Miller & Shepherd, 2009, p. 263). This multiplication made evident what was previously
hidden: the form did not define the social action, but rather the technology afforded a particular type of social action, only one of a number of possible social actions that it could and eventually did enable.

In this way, the emphasis on social action led them to conflate the genre with one element of its format, its distribution mechanism. As they explain, “The genre and the medium, the social action and its instrumentality, fit so well that they seemed coterminous, and it was thus easy to mistake the one for the other—as we did” (Miller & Shepherd, 2009, p. 283). This revisiting highlights the need to develop a terminology for identifying and discussing the various elements of format at play in the creation, sharing, and use of information resources. The formal characteristics of texts include ways of delivering texts and ways of creating meaning within texts that are not coterminous with the genres they enable.

Container Defined
As the example of blogs illustrates, one central consideration when discussing genres is how they are delivered to their intended audience. It is here that the concept of containers becomes meaningful. Containers are typified ways of collecting and presenting texts of particular genres for publication, understood in the broad sense of Hektor’s information activity framework which defines publishing as “posting or announcing formally or in public” (Gorichanaz, 2017b). Both of these activities are important: collecting for how texts are sought, chosen, and/or grouped; and presenting for how the collected texts are packaged and shared. In this way, containers are sets of practices and processes that are, at least in part, constitutive of how certain genres are enacted.

Social practices that shape publication emerge within communities in response to the need to share information beyond the groups in which people can personally interact. As such, they leverage the affordances of available communication technologies to meet rhetorical needs, and when they are successful in doing so, coalesce into recognizable publication practices. Sometimes, these practices transcend the communities that they emerged from and create systems of their own (e.g., the publishing industry), while other times they remain more localized (e.g., zines).

The practices of publication are co-constructive of social practices of information use. The outcomes of these practices and processes result in publications that are recognizable as instantiations of particular container types. Container types are recognizable because they coalesce as particular combinations of medium, mode, and technology that develop typified formal features. These typified formal features emerge from a combination of technological affordances, communicative needs, and the genres which the container publishes. Containers thus constitute an element of format that provides context for recognizing and responding to genres. The blog in Miller’s description above is actually a container type, a publication in which people share texts of the blog post genre type. What is worth noting here is that the container type helps to define the genre. What kind of post is it? A blog post. The same can be said of the widely used genre descriptor article. What comprises an article will vary widely depending on whether it is a news article, a magazine article, or a journal article.

The material instantiations of container types are a result of the social and material processes of their creation and the systems of knowledge and value in which they participate. As such, they provide crucial information that helps readers orient themselves to the genre and understand both its rhetorical moves and potentials. The example of the scholarly journal, as a container type that readers will be familiar with, is used here to illustrate these points.
The processes and practices by which texts are chosen for a container type or a specific publication delimit which genres are appropriate and what texts within a publication can mean in relation to one another. The scholarly journal container type, for example, gives strong preference to articles that report on research findings, but may also include a smaller number of book reviews, opinion pieces, or descriptions of emerging practices. Because most scholarly journals are organized around a particular discipline or subdiscipline, the texts chosen to be included in a given publication issue often do not have strong intertextual or thematic relationships with one another. However, it is expected that they will have those relationships with texts from other issues of the same journal or with texts from other parts of the field. This shapes the expectations of readers of these journals, orienting them as they seek information on specific topics or trace the evolution of particular conversations within their field.

Similarly, by establishing expectations for appropriate texts and maintaining processes of publication, container types shape how the texts within them are produced. The production of texts in this sense includes both the writing and publishing of those texts, which are interrelated but separate processes. Scholarly journals almost exclusively accept texts produced by scholars in the field represented by the journal, as an output of their research activities, and drawing on previously published research. Before publishing these texts, scholarly journals put them through an editorial process, usually including peer review, to determine their thematic relevance, scholarly rigor, and contribution to the field. Because of these processes of creation and publication, these texts are meant to be taken as an authoritative representation of the state of knowledge in the field while also being critically examined by other members of that field to be contradicted and/or built upon.

Container types emerge from and participate in particular social contexts and discourse communities, becoming a part of the structure of communication within these communities and governing to some extent the way knowledge can be created and shared. Scholarly journals emerge from scholarly research communities, specifically the disciplines and subdisciplines represented by the individual journal. These discourse communities are primarily composed of academic faculty who have a specific set of educational experiences, skills, and qualifications. The research and communication practices within these communities are highly codified, governed by both implicit community norms and explicit education and principles. These practices create sets of expectations for both writers and readers when engaging with the scholarly journal container type. In doing so, the rhetorical practices enabled by the container incorporate, replicate, and sometimes challenge the values and assumptions of those communities. By participating in the systems of power and prestige that privilege publication in certain containers and genres, both authors and readers can navigate their position within these discourse communities.

The container type also delimits the subject positions taken up by authors and readers. In the act of writing for the scholarly journal, authors prioritize their role as knowledge producers within their discipline over other available social positions and identities. In doing so, they prioritize certain ways of knowing over others, which will share some characteristics across disciplines but differ in others. For example, evidence and logical consistency are important elements of a scholarly argument. However, what counts as evidence will be different for microbiologists than it will for anthropologists. Taking up these subject positions and ways of knowing typically precludes the author’s other available subjectivities. For example, emotional appeals or arguments grounded in lived experience, though valid ways of knowing and communicating in many contexts, are generally discouraged in a scholarly argument.
As typified practices for the publication of information, container types also play important gate-keeping roles in the creation and sharing of information. Their role in structuring, and being structured by, communication and social engagement within a discourse community means that they are mechanisms of power that determine who gets to speak and be listened to, who is participating “correctly” and who is not. Genres and containers shape what is and is not sayable, what actions are and are not available. As sets of practices and expectations, they guide how people can create meaning and how people can interpret and respond. This does not preclude possibilities for creativity or resistance. However, creativity and resistance have to work both with and against genre and container expectations as ingrained patterns of behavior, as almost intuitive sets of understandings. As people internalize the expectations of genres and containers, their rhetorical purposes become entangled with the rhetorical, social, and pragmatic moves embodied by those expectations. From the outset, then, both genres and containers shape how people conceive of and create action through communication.

The emergence of the internet and digital communication technologies has created avenues for the publication and sharing of information that circumvent traditional publishing. Some predigital container types have made the shift to the new medium, with communities adapting the new technological affordances to their rhetorical needs and social practices. However, the new possibilities have also challenged the ways in which established container types structured the nature of communication within communities. While scholarly journals have, by and large, adapted to and flourished in the online communication environment, preprints have gained popularity as an alternative format for scholars to share texts that follow the expectations of the scholarly journal article genre outside of the publication practices of scholarly journals. Preprints have emerged as a result of changing values within the academic community, as historically closed models for sharing knowledge have been challenged and the community has begun to consider new models for more open sharing. It is important to remember that containers, and genres, are not fixed or immutable. Like languages, they change, often slowly and almost imperceptibly, as the values, needs, and practices of communities change and as changes to communication technologies afford new ways of creating and sharing information.

**Container and Genre Disambiguation**

Due to the close relationship between genres and containers, the two are often conflated. As demonstrated in Case Study 3 above, the unit of analysis has caused problems for LIS scholars working with rhetorical genre theory. Very few LIS studies of genre are explicit about the unit of analysis that they are trying to capture when using the concept. One exception is Santini, who states, “My unit of analysis is then the individual web page. Genres on the web can be also studied using other units of analysis, for example they can be analysed at website level…” (Santini, 2008, p. 703). Clarifying the unit of analysis through the disambiguation of genres and containers will help future research better identify genres and containers to ensure that each element of format can be consistently studied and compared.

When studying genres, the unit of analysis is the text. Texts can be subdivided, but their component parts are not meant to stand alone. In contrast, when studying containers, the unit of analysis is the publication. Their component parts are able to stand alone because those components are individual texts with potentially varying genre identities. A newspaper as a container type, for example, is made up of texts of varying genres, such as investigative reporting articles, opinion pieces, and classifieds, that can be read independently of one another.
Not all genres will necessarily have containers, and not all container types will necessarily contain multiple texts. The resume and cover letter, for example, are typically freestanding texts. In contrast, the scholarly monograph as a genre could be said to be coterminous with its container, a single text representing a single genre that takes up the entirety of a book. However, it is not uncommon for a scholarly monograph to have a foreword written by a separate author, or a dedication that is relatively unrelated to the rest of the content of the book. Genette (1997) has argued that it is these elements of peritext that turn the text into a book, suggesting that the content of the scholarly monograph genre can in fact be separated from the publication-resultant information that surrounds it and constitutes the book. However, the line between text and peritext is not as unambiguous as Genette suggests (Skare, 2020). While peritext is a useful way of thinking about how the publication process leaves behind artifacts of its practices, it may create artificial distinctions that do not always bear out in practice, particularly when considering the ways that genres of nonfictional prose are distributed outside of the publishing industry.

It is important to remember that genres are abstractions, not collections. Genres should not be understood as groups of texts created according to a specific set of genre conventions. Instead, they should be understood as mental models that people develop and deploy when trying to achieve certain types of genre actions. An individual text is an instantiation of a genre when others recognize it as such. Container types are also abstractions and an individual publication is an instantiation of a container type. However, part of the mental model for container types is what type of texts they collect and how they go about doing so. The newspaper container type, as described in the example above, is a mental model of how newspapers collect and publish texts, based on a history of interacting with various concrete instantiations of newspapers and an understanding of how news organizations produce them that may be derived through social interaction, direct experience, or research.

What this means is that genres and containers, despite being very closely related, are created according to different organizing principles and represent different potentials for rhetorical action. A genre represents a single action. It is created and organized according to the rhetorical or communicative purpose(s) that its creator(s) are trying to achieve, based on the conventions for action established by the genre’s form and the norms of the discourse community.

A container is usually compound. It is created according to some principle(s) of curation, selecting and combining texts based on its role within the community. A newspaper, for example, has the broad rhetorical purpose of keeping the community informed and includes texts by weighing considerations such as what the community should know (e.g., investigative reporting articles), what the community is interested in or needs (e.g., classifieds), and what represents the varying beliefs and values within the community (e.g., opinion pieces). Each text will have a rhetorical purpose of its own, more or less closely related to the purpose of the container. Newspaper articles that report on local events, for example, are very closely related to the newspaper’s purpose of keeping the community informed. An opinion piece written by a member of the community, however, may have a more complicated relationship to the newspaper’s purpose if it contains misinformation or disinformation. While the opinion piece does represent the community’s beliefs and values, it may run counter to the newspaper’s goal of providing reliable information.

Genres and containers also have different possibilities for completeness. Genres, as single rhetorical acts, are finished. Containers, on the other hand, may be continually accruing new texts as they
continue publishing according to their rhetorical principle. However, the era of digital publishing challenges the notion of static completeness. Digital texts can often be returned to and amended, sometimes without limit. This may contribute to the challenges of disambiguating these concepts in digital environments. While it is still helpful to consider genres primarily as completed texts, the fluidity of digital texts is an area that would benefit from further theory and study.

**Modes and File Types**

Although a full theoretical discussion of medium, modes, and multimodality is outside the scope of this paper, it is important to note that the framework for format created by the study team also borrowed concepts from the rhetorical study of multimodality to ensure the clearest possible disambiguation of genres and containers. In addition to container and genre as elements of format, the project team also distinguished file types and modes. Because the project took place in a simulated search engine environment, the medium of all of the information resources was the internet. However, participants were exposed to a variety of file types. File types were defined to include any file format, or standard for creating and storing information in digital files. This could include, for example, PDFs, HTML files, and ePub files.

Modes are “set[s] of socially and culturally shaped resources for making meaning” (Mavers & Gibson, 2012). Modes can be thought of as the building blocks for communication, or the tools that people use to create meaning. Written language has long been the dominant mode when studying information-seeking behavior. However, it is only one set of resources that people draw on to craft their messages. Images, charts and graphs, and other visualizations have become increasingly important ways of creating and sharing meaning, even in academic contexts. Video is another mode that has been made exponentially more common by the advent of the internet and the ease of creating and sharing information resources in that modality. Although these modes are sometimes confused for genres, the two elements of format are conceptually and functionally distinct. Texts in the same mode may belong to different genres, and similar genres may appear in different modes, such as a breaking news story that is published to the website as both a video recording and a written article. Most genres, especially online, incorporate multiple modes.

**An Example of Containers in Action**

To finish, it is important to see an example of these concepts applied to data to understand the ways in which they can help expand our understanding of information behavior. The following data from the RSIC project were originally reported on in Buhler et al. (2019) which provides a full description of the data collection methods. In that paper, Buhler et al. examine the effect of participant demographics on the likelihood that participants found three different information resources helpful, citable, and credible. The participants were students from four cohorts: high school, community college, undergraduate, and graduate. The three information resources examined all reported on the same scientific research study but represented different parts of the scientific communication lifecycle. The first resource was the journal article reporting the findings of the study in the Proceedings of the Royal Society B (RSPB), the second resource was a press release about the original journal article in the scholarly journal Nature, and the third was an article about the study findings in the popular magazine Time. Looking at these data broken down by container and genre will help to illustrate the role that each plays in participants’ evaluation of the resources.
Figure 2 shows the percentage of participants in each of the four cohorts who found each of the three resources helpful for the hypothetical school research project. Participants were not given a definition of helpful. The RSPB journal article was chosen as helpful most often across cohorts, the Nature press release was chosen as helpful least often across cohorts, and the Time magazine article was in the middle. In this instance, there is no pattern by container type. One resource in the scholarly journal was most helpful and the other was least helpful. However, if we break that down by genre, a pattern does emerge. Across cohorts, the articles were deemed more helpful than the press release.

**Figure 2**

*High School to Graduate School Helpful Judgments*

Figure 3 shows the percentage of participants who, having chosen each resource as helpful, also said that they would cite it in their hypothetical project. As in Figure 2, the total height of each bar represents the percentage of students in each cohort who chose it as helpful. The shaded portion of each bar represents the proportion of those students who also chose to cite it. Therefore, the smaller the clear portion is at the top of each bar, the more likely students were to cite the resource. Overall, students were most likely to cite the RSPB journal article, followed by the Nature press release, and least likely to cite the Time magazine article. In the citable judgments, we see the pattern from the previous helpful judgments reverse. There is no pattern by genre; one article is the most likely to be cited and the other is least likely. However, there is a pattern according to container type. Both resources in the scholarly journals are much more likely to be cited than the resource in the magazine, despite the magazine article being considered more helpful than the press release.

**Figure 3**

*High School to Graduate School Helpful and Citable Judgments*
Figure 4 shows the average credibility rating assigned to the three resources by each cohort. Participants were asked to rate the credibility on a scale of 1 (not credible) to 5 (highly credible) and were given a definition of credibility as whether the resource could be trusted and believed. The RSPB article was the most highly rated, with all four cohorts putting it in the mid-4 range. That was followed by the Nature press release, which had the largest variation across cohorts. On average, the lowest credibility scores were given to the Time article, with the exception of the community college cohort who rated the Nature press release lowest. The credibility judgments follow the same pattern as the citability judgments. The resources in the scholarly journals are considered overall more credible, while the resource in the magazine is considered less credible.

Figure 4

*High School to Graduate School Credibility Judgments*
These findings are not intended to be conclusive. Instead, they illustrate the analytic and explanatory power that can be gained by consistently applying the concept of genre and disambiguating it from container. What the above data suggest is that genre may be more important when students are determining helpfulness, but container may be more important when determining citability and credibility. Without clearly separating genre and container as elements of format, that finding would be obscured.

**Conclusion**

As typified ways of collecting and presenting texts of particular genres for publication, containers constitute an important element of format. They provide context for recognizing and responding to texts of various genres, helping people to navigate the complex information environment. While containers have formal features, they are first and foremost rhetorical and social. They emerge from and respond to the needs, values, and practices of discourse communities, evolving as the community evolves. Understanding containers requires attention to the processes by which information is formally and publicly shared. It also requires attention to the actions enabled by the genres that containers help to shape and distribute. The format of containers and the genres they contain are only meaningful when understood as outcomes of these processes and actions.

Distinguishing containers from genres, and situating them both in a framework with modes and file types, allows the RSIC project to develop a nuanced analysis of the role that each element of format plays in students’ point-of-selection behavior. This not only provides deeper and more precise insight into their information behavior, but also helps to inform the development of information literacy instruction grounded in students’ actual information-seeking practices. Additionally, rooting the study of format in information behavior in rhetorical theory eliminates formalistic approaches and foregrounds the social activity that makes formats meaningful. Doing so highlights the goals, values, and practices that guide individual information behavior within social and cultural contexts. This situates the study of
format in a more holistic understanding of information behavior. While the RSIC project focuses on students’ behaviors during a specific information activity in response to an imposed task, the framework of formats and the rhetorical perspective advanced in this paper can be applied to a variety of information activities across a variety of circumstances and environments. Doing so will help to build a nuanced understanding of the roles that formats play in information behavior that can be meaningfully compared across groups, activities, and contexts.

Acknowledgments
This project was made possible in part by the Institute of Museum and Library Services grant number LG-81-15-0155. The authors wish to acknowledge the research team members whose contributions made this work possible: Rachael Elrod, Randy A. Graff, Samuel R. Putnam, Erin M. Hood, Christine Towler, Robin Fowler, Summer Howland, Kailey Langer, Shakiyl Kirlew, Stefani Dopico, and Madelyn Wilson. We also wish to thank our advisory panel members for their insights: Alix Freck, Jennifer Kuntz, Jenna Miller, Megan Sorenson, Adam Fournier, Gayle Evans, Matthew Carrigan, and Emilio Bruna. The authors thank Julie Perino for her feedback on an early draft of this paper.
References


https://ufdc.ufl.edu/IR00010779/00001


https://doi.org/10.1080/02763915.2010.523327


https://multimodalityglossary.wordpress.com/medium/


https://doi.org/10.47989/irpaper881


https://doi.org/10.1080/10875301.2015.1082529


https://doi.org/10.1016/j.acalib.2018.07.007


Jankowski, A. (2018). “It was information based”: Student reasoning when distinguishing between scholarly and popular sources. *In the Library with the Lead Pipe.* http://www.inthelibrarywiththeleadpipe.org/2018/it-was-information-based/


http://www.ieee-tcdl.org/Bulletin/v2n1/rosso/rosso.html


https://doi.org/10.1016/j.ipm.2007.05.011


https://doi.org/10.1108/JD-01-2013-0002

Scott, R. E. (2016). Accommodating faculty requests and staying true to your pedagogical ideals in the one-shot information literacy session. *Communications in Information Literacy, 10*(2), 132–142.


https://muse.jhu.edu/article/185025


https://www.isko.org/cyclo/paratext


