Are Collections Sets?

Karen M. Wickett, Allen H. Renear
Center for Informatics Research in Science and Scholarship
Graduate School of Library and Information Science
University of Illinois at Urbana-Champaign
501 E. Daniel St
Champaign, IL 61820
wickett2@illinois.edu, renear@illinois.edu

Jonathan Furner
Graduate School of Education and Information Studies
University of California, Los Angeles
300 Young Dr N,
Mailbox 951520
Los Angeles, CA 90095-1520
furner@gseis.ucla.edu

ABSTRACT
The concept of a collection plays key roles in library, museum, and archival practice, and is arguably fundamental to information organization systems in general. Locating collection concepts in a reasonably robust ontology should have a number of practical advantages, including revealing inferencing opportunities and supporting consistency and coherence in system design and modeling. Although research practices involving collections have been studied empirically there has been surprisingly little attention given to the formal analysis of the concept itself, or to related notions like collection membership. With this paper we hope to convene that discussion, beginning with the question: Are collections sets? We consider in detail the substantial arguments against collections being a kind of set, but argue that at least one version of that claim cannot yet be ruled out. As ontology decisions, both practical and theoretical, are often more a matter of weighing competing considerations than making decisive arguments, progress on this question will probably require the comparison of alternative theories. We invite the information science community to join us in exploring the possibilities.

Keywords
Collections, modeling, ontologies, conceptual foundations.

INTRODUCTION
The concept of a collection is foundational to library, museum, and archival practice, and, we believe, fundamental to information organization systems in general. Collections of artifacts, images, texts, datasets, and other objects are not arbitrary aggregations, they are intended to support specific activities and are often carefully and deliberately designed to serve that purpose (Curral, Moss & Stuart, 2004; Lee, 2000, 2005; Palmer, 2004, 2006).

Many empirical studies in library and information science have addressed the role of particular kinds of collections in scholarly activities (Palmer, 2004) as well as the general features of collections as informational artifacts (Lee, 2000, 2005). However, there has been very little attention given to determining the ontological status of collections or the semantics of collection membership. It is not unusual to hear collections described variously as "sets" (Gonçalves, et al., 2004), "groups" (Galton, 2010), "aggregations" (CIDOC, 2010), or "selections" (Lagoze and Fielding, 1998). Yet it is unclear what exactly is intended by these terms when they are used casually, or even whether anything very specific and definite is intended at all.

Although we think that a notion this fundamental to information science deserves a systematic analysis for that reason alone, there are practical applications as well. The analysis presented in this paper grows out of an IMLS-funded project to develop inference rule categories for supporting reasoning from collection metadata (assertions about collections) to item metadata (assertions about items in collections) in RDF triple stores (Wickett, et al. 2010). The objective was to improve search, browsing, analysis, and information extraction in a registry containing both collection and item cultural heritage metadata. Similar applications are anticipated in the Data Conservancy, an NSF-funded project to develop cross-disciplinary repositories for scientific data curation.

While the category framework that was developed was largely determined by fairly obvious relationships between properties at different levels, we could see that there would be issues requiring a more precise understanding of the concept of a collection than we had available. It was apparent that how these uncertainties were resolved could make a difference to our emerging account of collection, shifting the pattern of inferences in one direction or another.

For both theoretical and practical reasons then we propose to convene an analysis of the concept of collection. In this paper however our focus will be on just one specific, but critical, question: are collections sets? More precisely the question is whether collections are sets (in the common mathematical sense of “set”) of their members.

Collections are certainly often described as sets. This is not only a frequent casual characterization (e.g. Lagoze and Fielding, 1998), but also found explicitly in formal models.
of digital library systems (Gonçalves, et al., 2008; Meghini, et al. 2010). This is not surprising. The colloquial notion of set seems to have ubiquitous application, the corresponding mathematical concept is well-defined and widely deployed, and upper-level formal ontologies frequently have sets as a fundamental kind of thing (e.g. Niles and Pease, 2001).

But are collections sets? And how do we tell?

**THE CONCEPT OF A COLLECTION**

The introduction of digital resources into library catalogs provided a valuable opportunity to review how collection development and management functions were addressed in libraries (Buckland, 1995; Atkinson, 1998). This was also an opportunity to “reconceptualize collection” (Lee, 2000; Casserly, 2002) in terms of how collections serve particular purposes for stakeholders, instead of relying on traditional notions of collections rooted in physical proximity. This shift in focus had a corresponding impact on evaluating user needs and how they can be met (Covi and Cragin, 2004; Kaczmarek, 2006).

For the purpose of developing a general conceptual account of collections, these reconceptualizations to accommodate digital collections provide deliberate and general characterizations of collections. Digitization projects continue to drive this work, both in the context of individual projects (e.g. M’kadem and Nieuwenhuysen (2010)) and in general (Lynch, 2002; Moss and Currall, 2004). However, these efforts have not, so far, resulted in a consistent accepted definition of collections. This gap has been noted several times since digital collections brought the issue into focus (Hill et al., 1999; Lee, 2000; Currall et al., 2004; Wickett et al., 2010).

In spite of the lack of clarity on what precisely collections are, arguments in favor of the usefulness of collection description point to benefits of collection description not for just management of collections, but for supporting scholarship and providing contextual information (Brack et al., 2000; Sweet and Thomas, 2000; Foulonneau et al., 2005; Palmer et al., 2006). Collection descriptions are designed to provide the contextual information (e.g. information about locations, times and related events, provenance, collection method) that aid scholars in making sense of the items within a collection.

Heaney (2000) provides an ER model of collections that has informed several schemas for collection description (Powell et al., 2000; Shrees and Cole, 2003; DCMI, 2007). In specifying what properties a collection can have and what relationships it can enter into such description schemas and, where they exist, conceptual models (like Heaney’s) provide insights into what collections are, as well as a useful starting point for terminology and informal definitions. These models also suggest examples of the sort of problems the lack of a robust definition of collection might cause when automatic inferencing and semantic technologies are common: an examination of collection level description reveals some attributes that are consistent with some possible concepts of collection and some with other concepts of collection.

However, the restricted expressiveness of ER and UML modeling languages, and the specific systems design focus of most of these conceptual models, limits their usefulness for our project here. Our analysis below will instead use first order logic, which provides additional expressiveness, precision, and clarity, and better supports identification of entailments. The meaning of each logical formula will also be expressed in ordinary English, so the reader unfamiliar with logical notation should not be at a disadvantage.

**THE ISGATHEREDINTO RELATIONSHIP**

We use the predicate $isGatheredInto(x,y)$ for the relation that stands between an item ($x$) and a collection ($y$) of which the item is a member. This name for the relationship comes from Heaney’s model of collections (Heaney, 2000), and is used in the Dublin Core Collections Application Profile (DCMI, 2007). We assume, until we have reason to believe otherwise that this is a distinct relationship that is not necessarily equivalent with any other known relationships from logic or ontology.

We use a one-place predicate $Collection(x)$ to represent the property of being a collection. The predicates $isGatheredInto(x,y)$ and $Collection(x)$ are clearly closely related, and we will discuss the possibility of reducing one of these concepts to the other later.

In the Dublin Core Collections Application Profile (DCMI, 2007), we have the following informal definitions:

- **Collection**: An aggregation of Items.
- **Item**: A physical or digital resource.

And the remark:

an Item is-Gathered-Into one or more Collections.

**Axioms relating Collection and isGatheredInto**

We now consider some possible axioms for $isGatheredInto(x,y)$.

The axiom below seems to follow immediately from the informal definitions of $isGatheredInto$ just mentioned – if something $y$ has something $x$ gathered into it, then that thing $y$ is a collection.

$$\forall y (\exists x isGatheredInto(x,y) \supset Collection(y))$$

On the other side of the $isGatheredInto$ relationship are items, which we can characterize with the predicate $Item(x)$, and another axiom that states that if something is gathered into a collection, then it is an item.

$$\forall x (\exists y isGatheredInto(x,y) \supset Item(x))$$

The converse of the first axiom states that if a resource is a collection, then it has something gathered into it.

$$\forall y (Collection(y) \supset \exists x isGatheredInto(x,y))$$
If we accept this axiom, then we in effect deny the existence of “empty collections”, collections with no members. Whether or not this is a reasonable axiom for isGatheredInto(x,y) is taken up later.

We can ask whether it is possible to gather a collection into a collection as a member by considering a domain restriction axiom.

\[ A4: \forall x \forall y (\text{isGatheredInto}(x,y) \supset \neg \text{Collection}(x)) \]

While there are often salient sub-collections within a collection, it is not clear that a collection that is part of a larger collection has been gathered into a collection in the same sense that items are. Given a case where collections are brought together to form a larger collection, we can make a distinction between understanding the curator to be gathering those collections into a larger collection and understanding the curator to be gathering the individual items from those collections into the larger collection on the basis of their membership in some previous collection.

**Relation Properties of isGatheredInto**

We can use our intuitions about the various relation properties of the isGatheredInto relation to assess further potential axioms with respect to collection membership. Since relation properties describe how a relation acts over a domain, we need to consider the domain for the formal characterizations of the relationship between collections and items. Previous work on collection/item metadata relationships (Wickett et al., 2010) developed rules that use a basic universal domain (the variables can take anything as values), and we take the same approach here. Cases that require further restriction of the domain can be handled by adding conditions to axioms as needed.

**Reflexivity**

A reflexive relation is one that every member of the domain stands in with itself. For example, greater-than-or-equal-to (≥) is a reflexive relation over the integers, since every integer is equal to itself.

A reflexivity axiom for isGatheredInto would be:

\[ R1 (\text{Ref}): \forall x (\text{isGatheredInto}(x,x)) \]

R1 states that everything is gathered into itself. Since that implies, by A1, that everything is a collection, R1 is clearly not a plausible axiom.

An irreflexive relation is one that no member of the domain stands in with itself. The relation greater than is irreflexive over the integers since no integer is greater than itself.

Irreflexivity of isGatheredInto would result in this axiom:

\[ R2 (\text{Irr}): \forall x \neg (\text{isGatheredInto}(x,x)) \]

R2 states that no collections are members of themselves. This axiom is much more plausible. It is not only difficult to imagine a collection that is a member of itself in a normal course of the use and development of collections of information resources, but it is hard to see what it would even mean for a collection to be gathered into itself.

**Symmetry**

A symmetric relation is one where, given any x and y in the domain, if x bears the relation to y, then y bears the relation to x. For example, the relation marriedTo is symmetric, since if x is married to y, then y is married to x.

Symmetry of isGatheredInto would give us the axiom:

\[ R3 (\text{Sym}): \forall x \forall y (\text{isGatheredInto}(x,y) \supset \text{isGatheredInto}(y,x)) \]

which would mean that every member of a collection has that collection as a member. This is certainly not plausible as a general axiom for collections.

An asymmetric relation is one where, given any x and y in the domain, if x bears the relation to y, then y does not bear the relation to x. Asymmetry of isGatheredInto would result in the following axiom:

\[ R4 (\text{Asym}): \forall x \forall y (\text{isGatheredInto}(x,y) \supset \neg \text{isGatheredInto}(y,x)) \]

which states that if something is a member of a collection, then that collection is not member of it in turn. Since it is difficult to imagine under what circumstances we would ever consider a collection to be gathered into one of its own members this seems a reasonable axiom for isGatheredInto.

An antisymmetric relation is one where, for any x and y, if x bears the relation to y and y bears the relation to x, then x and y must be the same thing. Greater-than-or-equal-to (≥) is an example of an antisymmetric relation.

\[ R5 (\text{AntiSym}): \forall x \forall y ((\text{isGatheredInto}(x,y) \& \text{isGatheredInto}(y,x)) \supset x = y) \]

If isGatheredInto is asymmetric, then asymmetry will follow as well, due to the “trivial” satisfaction of the conditional — the standard semantics for the material conditional (“⇒”) has the conditional true whenever the antecedent false. Since the conditional here will be true only because the asymmetry of isGatheredInto means that the antecedent of the conditional is always false, antisymmetry would not seem to give us a particularly useful axiom for our theory, though it is true nonetheless.

**Transitivity**

A transitive relation is one where, for any x, y and z, if x bears the relation to y and y bears the relation to z, then x bears the relation to z.

\[ R6 (\text{Trans}): \forall x \forall y \forall z ((\text{isGatheredInto}(x,y) \& \text{isGatheredInto}(y,z)) \supset \text{isGatheredInto}(x,z)) \]
If collections can be gathered into collections, and if collection membership is transitive, then whenever collection A is gathered into collection B, every member of A will also be a member of B (along with the collection A as an individual).

Transitivity again raises the question posed by A4: whether collections can themselves be gathered into collections. If it’s not possible for a collection to be an individual member of a collection, then transitivity follows, although only by trivial satisfaction of the conditional. At this point it’s most revealing to reject A4 and allow collections to be gathered into collections to see whether transitivity sensibly holds.

Allowing collections to be members of collections with transitivity would distinguish collections and isGatheredInto from sets and set relationships. Sets can be members of other sets, but set membership is not transitive: if set S is a member of set T, this does not mean that the members of S will themselves be members of T (although some or even all may happen to be).

While the subsetOf relationship between sets is transitive, it also has a distinct structure from the kind of transitivity for collections described above. If the set S is a subset of T, then every member of S is a member of T, but here S itself is not a member of T (although, again, it may happen to be).

If collections can be gathered into collections, as opposed to the items of some distinguished collections being gathered into a collection, then this creates a hierarchical structure within the collection. Allowing isGatheredInto to be transitive then collapses this structure. In order to preserve the intentions of curators who choose to gather whole collections instead of individual items from collections, we can consider transitivity not to hold for isGatheredInto. Non-transitivity will mean that isGatheredInto aligns with the set membership (memberOf) relation.

**Comparison of isGatheredInto with Other Relations**

There are some well-known relations that describe the structure of collective entities, most notably mereological (part/whole) relations and the set theoretic relations just mentioned. Examining how our intuitive understanding of isGatheredInto compares to the features of these relations can give us insight into collections. Table 1 shows a comparison between the relation properties of partOf, properPartOf, memberOf, subsetOf, and isGatheredInto.

The relation properties for partOf and properPartOf are based on the Classical Extensional Mereology as presented in Varzi (1996). The partOf relation is reflexive (everything is a part of itself), antisymmetric (if x is part of y and y is part of x, then x is y), and transitive (if x is part of y and y is part of z, then x is part of z). The properPartOf relation is irreflexive (nothing is a proper part of itself), asymmetric (if x is a proper part of y, then y is not a proper part of x), antisymmetric (as a trivial consequence of asymmetry), and transitive (if x is a proper part of y and y is a proper part of z, then x is a proper part of z). Neither of these merological relations match with isGatheredInto, which is irreflexive and asymmetric (like properPartOf) but not transitive.

The relations properties for memberOf and subsetOf are based on a standard Zermelo-Fraenkel axiomatization of set theory (ZFC) (Fraenkel and Bar-Hillel, 1958). The set membership relation memberOf is irreflexive (no set can be a member of itself), asymmetric (if x is a member of set y then y cannot be a member of x), antisymmetric (again, a trivial consequence of asymmetry) and not transitive. The subsetOf relation is reflexive (every set is a subset of itself), anti-symmetric (if x is a subset of y and y is a subset of x then x is y) and transitive (if x is a subset of y and y is a subset of z, then x is a subset of z).

As Table 1 shows, the relation properties of isGatheredInto align with the relation properties of the set theoretic relation memberOf. Although this result supports the claim that collections are sets, it does not automatically provide a decisive answer to our question. As we argue later in this paper, it is still difficult to reconcile all of the properties we commonly assign to collections with the properties of sets.

**Table 1: Comparison of relation properties**

<table>
<thead>
<tr>
<th>+/-</th>
<th>partOf</th>
<th>properPartOf</th>
<th>memberOf</th>
<th>subsetOf</th>
<th>isGatheredInto</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ref</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Irr</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Sym</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Asym</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>AntiSym</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Trans</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
</tbody>
</table>

**D1:** \( \forall x (\text{Collection}(x) \implies \exists y \text{ isGatheredInto}(y, x)) \)

The notation "=\( \exists y \)" is read "if and only if" as definitions of this sort are typically understood as implying logical equivalence.

Because D1 is just the joint assertion of A1 and A3 it has the consequence, (from A3) that there are no empty collections. Can we be confident this is so? It is easy to imagine cases where curators deem a collection to have been created (in terms of allocating resources, or giving a collection a name or preliminary description) without having, at that time, gathered any items into the collection.
In order to handle these cases we can modify the definition to say that something is a collection if and only if it is possible that there is something gathered into it.

D2: \( \forall x (\text{Collection}(x) = \exists y \exists z \text{isGatheredInto}(y,z,x)) \)

D2 is read “for all \( x \), \( x \) is a collection if and only if it is (logically) possible that there exists a \( y \) such that \( x \) is gathered into \( y \).” Or alternatively, “... if and only if \( x \) could have something gathered into it” This definition uses the modal operator for possibility (“\( \exists \)”), and so requires the use of a more expressive language than first order logic. Or one might rather forgo empty collections and say that while curators may have created something when they designate resources for the development of a future collection, they have not (yet) created a collection.

Of course even if either of these definitions is accepted, our understanding of collections is not advanced much. The concepts of collection and isGatheredInto are so closely related that any definition of one in terms of the other is unlikely to provide many satisfying analytical insights.

In addition these definitions do not, as they stand, give us identity conditions for collections. That is, while such a definition will tell us whether or not something is a collection, they do not provide a way to distinguish different collection, or to identify a specific collection as continuing to exist despite changes over time, and that is a critical part of providing an account of the ontological nature of collections. In the next section we consider whether or not collections are sets, and here a positive answer will in fact provide identity conditions.

**COLLECTIONS AS SETS**

Our table shows that isGatheredInto appears to have the same relation properties as the set membership relation: both are irreflexive, asymmetric, and non-transitive. And not surprisingly sets are indeed commonly used to represent collections in models of digital library systems (Lagoze and Fielding, 1998; Gonçalves et al., 2004; Meghini and Spyratos, 2010). However, it is not always clear if these models are treating collections and collection as equivalent notions, or if collections are being represented as a kind of set. If the former then all collections are sets and all sets are collections. But if the latter then while all collections are sets, not all sets are collections.

**Collections as equivalent to sets**

In a formal logic-based model for digital libraries (Gonçalves et al. 2004), we have the definition:

A collection \( C = \{d_{o1}, d_{o2}, ..., d_{ok}\} \) is a set of digital objects.

Although this definition does limit collections to a kind of set (sets of digital objects), it seems to imply that any arbitrary set of digital objects is a collection, regardless of whether that set has received any sort of recognition or attention from a curatorial agent. That is, regardless of whether or not any collecting has occurred.

Lagoze and Fielding (1998) define a collection as “a set of criteria for selecting resources from the broader information space.” Taken literally, this phrasing implies that a collection is a set of criteria, which seems peculiar. But this account can be read more charitably as intending that any set of resources meeting a set of criteria is a collection (“set membership ... is ... criteria-based”).

These definitions leave us with the same basic problem. In Gonçalves et al. (2004), all sets of digital objects, however arbitrary, are considered collections, and since every set of resources meets some set of criteria or other, the same is true with the account in Lagoze and Fielding (1998). We make this argument in more detail below.

**Collections as a kind of set**

The addition of further restrictions (e.g. “a set identified by curators”) will get us a more intuitive view, with some but not all sets being collections, although the identification and precise formulation of such a criterion will not be easy.

However, even if we identify a criterion that seems to cut the cases correctly, giving us only, and all, the sets that are collections, a separate issue will pose a challenge regardless of what criterion is chosen. It is natural to say that collections can change their membership — items are often removed from collections, and collections are often expanded to include new topics or creators. Sets, on the other hand, cannot lose or gain members. So how could collections be sets if collections can have properties (becoming larger or smaller) that no set has?

**ARGUMENTS AGAINST COLLECTIONS AS SETS**

We now present in detail the two arguments against collections being sets that were mentioned briefly above. The exposition will be methodical and detailed not just because the belief that a collection is a set is widespread, but also because there seem to be few good alternatives to collections being sets, and doing without sets as collections will almost certainly pose substantial modeling and ontology challenges. In addition, although the arguments here may seem too simple to warrant such methodical treatment, Sharvey (1968) and van Cleve (1985) have shown that in this area simple and apparently valid arguments have received wide endorsement before their fallacies were identified.

We present counterexamples against both of the claims mentioned above: (i) that the collection and set are equivalent concepts, and (ii) that a collection is a kind of set. We begin by replacing each of these two claims with weakened versions that are entailed by the corresponding originals. We then argue against the weaker version directly. Since the weaker versions are entailed by the originals any successful arguments against the weaker versions will also be effective against the original claims.

We begin with:

**E1:** Collection and set are equivalent concepts
Understanding what claim is being made by \( E1 \) is complicated by the several possible senses of "equivalent", ranging from strict identity of both meaning and extension ("collection" and "set" have the same meaning and apply to the same things; i.e. they are synonyms), to some sort of definitional or analytical equivalence, to logical equivalence, to material equivalence. Although material equivalence (all collections are sets and all sets are collections, but just as a matter of fact, at the moment of assertion) may seem too weak to be the intended sense of \( E1 \), it is logically implied by any relevant variety of equivalence that might be the intended sense of \( E1 \). Consequently if this weak sense of equivalence, material equivalence, fails — because either some collections are not sets, or some sets are not collections — then there is no relevant sense of (logical) equivalence in which collections and sets are equivalent.

The weakened form of \( E1 \), then, is:

\[
E2: \forall x (\text{Collection}(x) \supset \text{Set}(x)) \& \forall x (\text{Set}(x) \supset \text{Collection}(x))
\]

If something is a collection then it is a set and if something is a set then it is a collection.

**Against collection and set being equivalent concepts**

Our criticism of \( E1 \) takes the form of counterexamples to \( E2 \). If those are successful then \( E1 \) is false, regardless of the intended equivalence. Our argument against \( E2 \) will take the form of counterexamples against its second conjunct:

\[
E3: \forall x (\text{Set}(x) \supset \text{Collection}(x))
\]

If something is a set then it is a collection.

A counterexample against \( E3 \) will be a case where something is a set but not a collection.

\( E3 \) does indeed seem to have serious counter-intuitive consequences when combined with any standard axiomatization of mathematical set theory. The most important of these would have collections existing without intentional activity on the part of curatorial agents. In any set theory with “unrestricted comprehension” there is, for any two things in the world, a set that has just those two things as members. So there is, for instance, a set that contains the planets Mars and the tallest redwood tree in California. There is such a set, but is there such a collection? If we think not, if we believe that some sort of curatorial agency is required for the existence of collections over and above the existence of sets, then this is a reason to reject \( E3 \) and conclude that not all sets are collections.

It might be argued that once Mars and the redwood tree have been indicated by someone (as we have indicated it just now) then the collection does indeed exist — and so this is not a case where a set is not collection. However the set in question also existed ten years ago (by, again, the unrestricted axiom of comprehension), well before we, or anyone, indicated that set. If ten years ago that set existed but was not a collection then not all sets are collections.

Naive set theory, with unrestricted comprehension, is of course prey to Russell's paradox (Is the set of all sets that are not elements of themselves an element of itself? If it is then it isn't and if it isn't then it is). The set theories that avoid Russell's paradox typically define new sets in terms of existing sets and so for those axiomatizations the general form of the argument just given is not available (Fraenkel and Bar-Hillel, 1958). Nevertheless it is routine in mathematics and ontology to assume that for any two objects there is a set that contains them both and that assumption, at least when restricted to objects other than sets, is consistent with ZFC and does not generate Russell's paradox — and does provide a counterexample to \( E3 \).

In any case it is a consequence of ZFC, which does not allow unrestricted comprehension, that if one set exists then an infinite number of sets exist (by repeated applications of the axiom of pairing). But it would be wildly at odds with our ordinary notion of collection to say that if one collection exists then there are an infinite number of collections — few if any of which having received any curatorial attention whatsoever.

One might argue that implications of this latter sort are consequences of axiomatizations that do not match our pre-theoretical notion of a set, and that an improved axiomatization could be developed, one that reflected what real sets (sets for ordinary people, not the sets of mathematicians) really are. But it is hard to see how any axiomatization could simultaneously capture most of the common assumptions about mathematical sets (for instance, that for any two things there is a set that has just those two things as members) and yet not have counterintuitive consequences for the assertion that all sets are collections.

**Against collections as a kind of set**

The second sense of an affirmative answer to "Are collections sets?" is:

\[
K1: \text{A collection is a kind of set.}
\]

Whereas the claim that sets and collections are equivalent may have seemed implausible on its face (implying as it does that all sets are collections), the claim that a collection is a kind of set (which allows that some sets are not collections) is quite common, and appears in LIS literature and formal models of digital libraries, as we have shown.

Although it might be difficult to say exactly what the "is a kind of" relationship amounts to, by any reasonable construal it implies at least that all collections are sets, and specifically it implies the material conditional form of that assertion. And so our weakened version of \( K1 \) is:

\[
K2: \forall x (\text{Collection}(x) \supset \text{Set}(x))
\]

If something is a collection then it is a set.
We will consider arguments that \( K2 \) is false and that therefore a collection is not a kind of set. As with \( E1 \) our criticism of \( K1 \) will take the form of counterexamples against \( K2 \). If those counterexamples are successful then they are counterexamples against \( K1 \) as well, regardless of the precise meaning of "is a kind of".

Since \( E2 \) also implies \( K2 \), any arguments against \( K2 \) will also be arguments against \( E2 \), and therefore arguments against \( E1 \). That is, our evidence against "a kind of" is also evidence against equivalence.

We first note that for every collection there exists a set that has as members all and only those things gathered into that collection.

\[
S1: \forall x \forall y \forall t (Collection(x,t) \supset \exists z (memberOf(y,z,t) = isGatheredInto(y,x,t)))
\]

\( S1 \) asserts that for every collection \( x \) at a time \( t \) there is a set \( z \) that has as its members all, and only, the items in the collection. We use "\( memberOf(x,y) \)" here only in the strict sense of set membership and introduce time indices.

If a collection is a set, what set might it be? Obviously one candidate is the set just identified, the set of all things that are items in that collection. Certainly that set exists, and certainly the set and the collection have a lot in common (their items/members for one thing).

That the set of any collection’s items exists is not in dispute. Our question is not whether there is some set that has the same members as \( C \), but whether the collection \( C \) is that set, that is, whether \( S \) and \( C \) are identical. The claim that a collection and the set of its items are identical can be made by adding "\( x = z \)" to the biconditional in \( S1 \).

\[
S2: \forall x \forall y \forall t (Collection(x,t) \supset \exists z (memberOf(y,z,t) = isGatheredInto(y,x,t) \& x = z))
\]

This is the claim we are considering now. We argue that it is inconsistent with common beliefs about sets and collections. Specifically it is inconsistent with affirming Member Essentialism (van Cleve, 1985) for sets while denying Item Essentialism for collections.

Member Essentialism can be expressed this way:

\[
ME: \forall x \forall y \forall t ([\exists (y,t) \& memberOf(x,y,t)] \supset \forall x (exists(x,t) \supset (exists(x,t) \& memberOf(x,y,t))])
\]

Member Essentialism says that if some set \( y \) has \( x \) as member at some time \( t \), then \( y \) has \( x \) as a member, necessarily that is, \( y \) has \( x \) as a member at every time \( x \) exists and in any possible alternative circumstance (or "possible world").

Member essentialism is widely accepted by mathematicians, philosophers, and ontologists. The basic idea is that a set, e.g. \( \{a,b,c\} \), could not have had in the past, or have in the future, or have in some alternative circumstance (any other possible world) any other members than those it actually has, which are, in our example: \( a,b, \) and \( c \). Consider the temporal case first. How can set \( \{a,b,c\} \), for instance, come to have, at some point in the future, only two members, \( a \) and \( c \), and so become the set \( \{a,c\} \)? Exactly what would the persisting entity be that at one time had \( a,b, \) and \( c \) as its (only) members and then at a later time had just \( a \) and \( c \) as members?

Since only sets have members only a set will do for the thing that once had \( b \) as a member and then later does not. But exactly what set would it be? Obviously it can’t, on pain of immediate contradiction, be the set \( \{a,b,c\} \), as that set does have \( b \) as a member and so doesn’t meet the criterion of not now having \( b \) as a member. So we are left with \( \{a,c\} \) as a candidate for the set that once had \( b \) as a member and now does not. It is indeed true that \( \{a,c\} \) does not now have \( b \) as a member, so that half of the requirement is satisfied. But is it really the case that the set \( \{a,c\} \) might at some point have had \( b \) as a member?

It is true that we say things like "if the value 42 is removed from the set \( S \) the result of applying the formula will be different", but this seems just a manner of speaking. We may say that we removed \( b \) from the set \( \{a,b,c\} \), and even that the numerical size of some set changed (in this case decreasing by one). But it seems more accurate to make the same point by saying instead that we turned our attention from one set, a set with three members \( (a,b, \) and \( c) \), to another set, one with just two members \( (a \) and \( c) \).

The claim that sets cannot lose or gain members is a claim specifically about sets and their members. It is not deduced from the (implausible) claim that nothing can lose or gain a property, nor does it imply that thesis. There is no difficulty in assuming that a person can be happy at one time and not happy at a later time, or that a leaf can be green at one time and not green at later time. In both cases we easily identify an underlying enduring object. Nor do we assume that Member Essentialism for sets implies or is implied by mereological essentialism, the thesis that composite physical objects have their parts essentially.

In ordinary discourse we also say things like “the set of people living at 303 Main St is larger than it was used to be.” To make the problem acute let’s name that set \( F \) and define it using by a rule, using “set builder” notation:

\[
B1: \quad F = \{x: \text{livesAt303Main}(x)\}.
\]

On Monday just John and Jill live at 303 Main; so let

\[
M = \{\text{John}, \text{Jill}\}.
\]

On Tuesday John, Jill, and Mike live at 303 Main, so let

\[
T = \{\text{John}, \text{Jill}, \text{Mike}\}.
\]

Is the set \( F \) larger on Tuesday than it was on Monday? No. There are two ways to understand what set “\( F \)” refers to at a given time, but in neither case does the set \( F \) “get larger”.

1) If \( B1 \) is understood as assigning to “\( F \)” the set of objects that satisfy \( \text{livesAt303Main}(x) \) at the time \( B1 \) is asserted, and \( B1 \) is asserted on, Monday, then on Monday “\( F \)” refers
to M, i.e. \{John, Jill\} and on Tuesday “F” continues to refer to M, \{John, Jill\}. By this convention “F” refers to the same two person set on two different days, and not to some single set that got larger between Monday and Tuesday.

2) Alternatively B1 may be understood to define “F” as referring to whatever set satisfies livesAt303Main(x), at the time “F” is used. In that case “F” refers on Monday to set M \{John, Jill\} and on Tuesday to set T \{John, Jill, Mike\}. But since these are two different sets there is not, again, any set that once had two members and then later had three.

Since Member Essentialism for sets is believed by many to follow directly from the ZFC axiom of extensionality it may seem we belabor this point too much. However Sharvey (1968) and van Cleve (1985) have shown that ME does not follow from that ZFC axiom (alone) and in fact its basis remains obscure. The arguments just given are intended to reiterate and confirm the intuitive plausibility of Member Essentialism, not decisively deduce it from other principles. Fortunately, why ME is true need not concern us here — it is almost universally affirmed and does indeed appear to be part of our understanding of sets.

Now let’s consider a corresponding principle for collections, Item Essentialism.

\[
\text{IE: } \forall x \forall y \forall t_1 (\text{Collection}(y, t_1) \land \text{isGatheredInto}(x, y, t_1)) \supset \\exists \forall t_2 (\text{exists}(y, t_2) \supset (\exists x (\text{exists}(x, t_2) \land \text{isGatheredInto}(x, y, t_2)))
\]

Item Essentialism says that if some collection \(y\) has something \(x\) as an item at some time \(t_1\), then that collection has \(x\) as an item whenever \(x\) exists and in every possible alternative circumstance.

While ME is universally accepted as a necessary truth about sets, IE seems to conflict with our settled conviction that (i) items may be added to collections and removed from collections and (ii) collections could have had items other than the items they do have — imagine a failed attempt to acquire an item: if the attempt had succeeded the collection would have had an item it does not have.

In summary, if collections are sets, with the items gathered into them as the members of those sets, then Item Essentialism would be part of our concept of what a collection is. But if Item Essentialism is not part of our concept of what a collection is, then collections are not sets.

There are just two ways out: deny ME for sets or affirm IE for collections. But denying ME for sets would seem to be out of the question, so that leaves affirming IE for collections, that is, holding that collections cannot add or gain members. That seems a high price to pay, conceptually, for classifying collections as a kind of set.

**COLLECTION AS SET-IN-A-ROLE**

Understanding collections to be a kind of set was initially promising and natural. But if collections aren’t a kind of set, then what are they? It is not likely that the answer to this question will be easy either to develop or to defend.

Moreover, while we have seen that there are substantial counterintuitive consequences if collections are understood as a kind of set, perhaps the counterintuitive consequences of the alternative treatments are worse. Before we bring to a close this first installment in our consideration of the ontological status of collections, we want to make the strongest case possible for collections being sets of their members. This will be a case that accepts Item Essentialism as an unavoidable, though counterintuitive, consequence.

This approach to the ontological status of collections holds that a collection is a set in a particular informational or curatorial role. Being a collection is a property that sets have only in certain contingent circumstances. So on this account, sets that have not received any kind of curatorial attention (e.g. the set that contains the planet Mars and the tallest redwood tree in California) exist, but do not qualify as collections. Although a set exists whenever its members exist, a set is not a collection unless it is treated as such in the appropriate social circumstances. Therefore sets that aren’t collections can become collections, but nothing that is not a set can be collection.

On this account the property of being a collection is what Guarino and Welty (2000) refer to as a non-rigid property. Guarino and Welty define rigidity using modal logic and model theory (the notion is based on the idea of de re necessity used in ME and IE), but the basic idea is simple: a property is rigid if and only if nothing that has that property could have both (i) existed and (ii) failed to have that property. For example, being a person is rigid because the things that are persons could not have been anything but persons, but being a student is not rigid because the things (persons) that are students might not have been students.

Becoming a student or ceasing to be a student (without ceasing to exist entirely) is only possible if being a student is a non-rigid property. According to Guarino and Welty rigid properties indicate types, fundamental kinds of things, while non-rigid properties indicate roles that things of some particular type may enter into.

On this account the property of being a collection is what Guarino and Welty (2000) refer to as a non-rigid property. Guarino and Welty define rigidity using modal logic and model theory (the notion is based on the idea of de re necessity used in ME and IE), but the basic idea is simple: a property is rigid if and only if nothing that has that property could have both (i) existed and (ii) failed to have that property. For example, being a person is rigid because the things that are persons could not have been anything but persons, but being a student is not rigid because the things (persons) that are students might not have been students. Becoming a student or ceasing to be a student (without ceasing to exist entirely) is only possible if being a student is a non-rigid property. According to Guarino and Welty rigid properties indicate types, fundamental kinds of things, while non-rigid properties indicate roles that things of some particular type may enter into.

On the set-in-a-role view a collection is not a type of thing, but a role that things of some type or other have in particular circumstances. What sort of thing is it that can enter into the role of a being a collection? *Sets*. Being a collection is a role that sets have in the right contingent social circumstances. That being a collection is a role is suggested by the fact that being a collection is not rigid: the thing (the set) that is a collection might not have been a collection — it might not have received any curatorial attention. On the other hand the property of being a set *is* rigid: nothing that is a set could have been anything other than a set, and things that are sets in any possible world are sets in every possible world. Being a set then is a type and therefore a candidate for bearing a role.

On this account of collections, collections are sets, and so they have the identity conditions of sets: they cannot lose or gain items. This is the counterintuitive consequence we had
indicated would be the price of affirming Item Essentialism in order to have collections be sets. But there is some consolation in noting that there are alternative ways to characterize the fact that collections can lose and gain members. When we say that a collection has undergone a change in membership, we mean that a different set has been selected for the purpose at hand by some particular person or persons. The expectations and practices for the development, management, and description of collections, in combination with social conventions and practices, create a system for recording and communicating the set that is currently distinguished as the relevant collection.

Even though this approach does imply that items cannot be literally added to or removed from persisting collections, it does not require treating sentences such as “We added an item to the collection” as false. Such sentences may be considered idioms that while not literally true, are true when given the appropriate elucidation. This strategy for defining collection and changes in collection size is similar to one proposed for defining document and changes in documents (Renear and Wickett, 2009). A remaining challenge will be to provide an account of cases where two sets are considered the “same collection” — if they cannot literally be the same collection.

STATE OF PLAY
We have considered the arguments against collections being sets, and noted that any theory of collections as a kind of set will contradict the widely held belief that collections can lose or gain members, as well as the related belief that collections might have had members other than the ones they do have. That may seem like a decisive refutation. However until it has been shown that there is an account that will accommodate all of our pre-analytic beliefs about collections we must allow that defining collection may be less a matter of decisive argument or counterexample, than the weighing of pros and cons of different positions, once those positions have been developed. Revised notions of membership change and collection identity may not be such high price to pay for avoiding what might be regarded by some as even more unpleasant alternatives. These alternatives range from fundamentally new ontological kinds and relationships to simply denying the existence of collections altogether:

1) *Sui generis* ontological entities.

2) Sets, but not sets of their members (Effingham, 2010).

3) Things “constituted” by sets, where constitution is either (i) a unique unanalyzed relationship (Baker, 2007); or (ii) defined in terms of things and properties already accepted in the ontology (Uzquiano, 2004).

4) Things that temporally “overlap” with their members (Korman, 2010).

5) Things synchronically identical to a mereological fusion of their members (Galton, 2010; CIDOC 2010).

6) Predicates, with intension and extension (Meghini and Spyropoulos, 2010).

As we said above, we are convening a discussion and not, at this point, arguing for any particular conclusion. We invite the information science community to join us in exploring the possibilities.

ACKNOWLEDGEMENTS
We would like to acknowledge the UIUC/GSLIS Metadata Roundtable (2003-2010) for theoretical discussions on the nature of collections, and, particularly Carole Palmer, Muriel Foulonneau, Tim Cole, and Richard Urban. This work has been partially supported by an IMLS grant (LG-06-07-0032-07) and NSF grant OCI-0830976.

REFERENCES


CIDOC CRM Special Interest Group (2010). Definition of the CIDOC Conceptual Reference Model, version 5.0.2. [http://www.cidoc-crm.org/docs/cidoc.crm_version_5.0.2.pdf](http://www.cidoc-crm.org/docs/cidoc.crm_version_5.0.2.pdf)


