The Life Cycles of Genres Ted Underwood[1]

The concept of genre is as old as literary theory itself, but centuries of debate haven't produced much consensus on the topic. Part of the reason is that genre looks like a different thing at different points in the life of a text. Scholars of rhetoric tend to focus on the patterns of communicative action that produce memoranda or tragedies (Devitt 2004; Miller 1984). Sociologists are sometimes more interested in institutions that organize reception (Bourdieu 1984; DiMaggio 1987). Literary scholars, for their part, have traditionally been preoccupied with the patterning of the texts themselves. Of course, all of these aspects of genre are connected. But it's not easy to describe the connections.

Distant reading may seem to lend itself, inevitably, to literary scholars' fixation on genre as an attribute of textual artefacts. But the real value of quantitative methods could be that they allow scholars to coordinate textual and social approaches to genre. This essay will draw one tentative connection of that kind. It approaches genre initially as a question about the history of reception — gathering lists of titles that were grouped by particular readers or institutions at particular historical moments. But it also looks beyond those titles to the texts themselves. Contemporary practices of statistical modeling allow us to put different groups of texts into dialogue with each other, in order to discover, for instance, whether competing definitions of the Gothic (created at different times and embodied in entirely different lists of works) were nevertheless as compatible as some critics claim.

[1] The author gratefully acknowledges the support of the Novel™ project, funded by Canada's Social Sciences and Humanities Research Council. Support for this project was also provided by Knowledge Lab at the University of Chicago, and by the Center for Advanced Study at the University of Illinois, Urbana-Champaign. HathiTrust Research Center guided my exploration of HathiTrust Digital Library; useful advice came from all the members of the Novel™ team, and from James English.

The problem of historical comparison is a pressing one because literary scholars haven't been able to reach much consensus about the life cycles of novelistic genres. The Gothic, for instance, can be treated as a category that lasts for 25 years or for 250. In *Graphs, Maps, Trees*, Franco Moretti surveyed academic studies of genre, and concluded that genres display "a rather regular changing of the guard . . . where half a dozen genres quickly leave the scene, as many move in, and then remain in place for twenty-five years or so" (Moretti 2005: 18). The Gothic novel of the late eighteenth century gives way (say) to the Newgate novel around 1825, then to the sensation novel in the (late) 1850s, and eventually to the late-nineteenth-century "imperial Gothic" (e.g. *Dracula*), which Moretti's chart treats as a phenomenon entirely different from the older Gothic of monks and banditti. Moretti speculatively links the twenty-five-year rhythm of this sequence to generational succession.

On the other hand, there are well-established traditions of reception that suggest genres can sustain a coherent identity over much longer timelines than this generational rhythm would allow. Fans of contemporary mystery fiction often read Agatha Christie and Wilkie Collins with equal pleasure (James 2011). And critics, at least, enjoy framing Stephen King as the inheritor of a Gothic tradition that stretches back continuously through H. P. Lovecraft and Bram Stoker, all the way to *The Castle of Otranto* (Sears 2011: 51, 174).

We know that all these claims are valid as statements about reception. Moretti is right that many academic studies of genre do cover generation-sized periods. But it is also true that categories like detective fiction have mattered continuously to readers for more than a century. Textual analysis won't prove either claim wrong, but it may help us understand how they're compatible. For instance, one obvious way to reconcile conflicting accounts might be to say that Moretti is right about the rhythms of genre in the century he discusses (the nineteenth) — but wrong about the twentieth, because genres harden there into durable marketing institutions. Moretti, however, has hinted (2005: 31) that even a

long-lived twentieth-century genre like detective fiction might be at bottom a sequence of generational stages (the Holmesian "case", the closed circle of country-house suspects, the crime thriller), linked by a relatively weak thread (see Roberts 2011 for an analogous theory of science fiction). The beginning of a sequence like this might not even resemble the end. Maurizio Ascari, for instance, argues that "the term *detective fiction* has been increasingly supplanted by *crime fiction,*" and glances skeptically at the old narrative that positioned "Poe as a founding father" of the whole tradition (Ascari 2014: 15).

In short, critics like Moretti and Ascari argue that the social continuity of readers' interest in detective stories has veiled a series of formal discontinuities. This is where textual evidence could start to provide a useful test. Matthew Jockers has shown that genres framed on a twenty-five- to thirty-year scale are linguistically coherent phenomena in the nineteenth century. A statistical model trained on examples of silver-fork or sensation fiction can identify other examples of the same genre with reasonably high accuracy (Jockers 2013: 67-81).

It would be interesting to discover whether this works equally well for books linked by a longer-lived tradition, like the detective story. Intuitively, one might expect a century-long group to be looser; it's hard to believe that *The Moonstone* (1868) and *The Big Sleep* (1939) really have much in common stylistically. If it does turn out to be easier to recognize sensation novels than detective fiction, we'll have some evidence that Moretti was right about the underlying generational logic of genre. This evidence wouldn't rule out the possibility of longer-term continuity: we don't know, after all, that books need to resemble each other textually in order to belong to the same genre. But we might conclude at least that generation-sized genres have a particular *kind* of coherence absent from longer-lived ones.

On the other hand, we might just as reasonably expect to find a very different historical pattern. Scholars have spent a great deal of energy tracing the gradual standardization of genre conventions in the early twentieth century, pointing to genre-specific pulps and critical pronouncements like Ronald Knox's

so-called "Decalogue" (1929) of rules for detective stories as moments of genre consolidation (Knox 1976). For some genres the process is thought to have taken even longer. Gary K. Wolfe suggests that "the science fiction novel persistently failed to cohere as a genre" until the 1940s (Wolfe 2011: 21). If this account of literary history is correct, we wouldn't expect to find a succession of distinct generational phases, but a steady hardening of boundaries, producing genres that are much more clearly distinct by the middle of the twentieth century than at its outset. That's the story I expected to find when I began this project.

To investigate these questions, I've gathered lists of titles assigned to a genre in eighteen different sites of reception. Some of these lists reflect recent scholarly opinion, some were defined by writers or editors earlier in the twentieth century, others reflect the practices of many different library catalogers (see Appendix A). Although each list defines its object slightly differently, they can be loosely arranged around three categories: detective fiction (or "mystery" or "crime fiction"), science fiction (also defined in a variety of ways), and the Gothic. (It is debatable whether the Gothic, writ large, is a genre at all — but that's what makes it an interesting case.) I also collected texts corresponding to these titles, relying on the Chicago Text Lab and HathiTrust Digital Library as sources.[2] By comparing groups of texts associated with different sites of reception and segments of the timeline, we can ask exactly how stable different categories have been.

The story that emerged from this experiment doesn't line up very neatly with either of the alternative accounts I just gave: generational succession or gradual consolidation. I see little evidence of the generational waves Moretti's theory would predict. In fact, it's not even the case that books in a chronologically

[2] The process of normalizing texts hides pitfalls for the unwary. For instance, I have tried to remove front matter, prefaces, and running headers, which otherwise might include explicit genre labels the model would seize on. It is also potentially a concern that these texts are drawn from two different libraries (HathiTrust and Chicago Text Lab). But in practice, I don't observe the chronological break one might expect where the two sources join (in 1923). Models trained on one library predict genres in the other without difficulty.

focused genre (like "the sensation novel, 1860-1880") necessarily resemble each other more closely than books spread out across a long timeline. Detective fiction and science fiction display a textual coherence that is at least as strong as Moretti's shorter-lived genres, and they sustain it over very long periods (160 or perhaps even 200 years). So I think we can set aside the (productive) conjecture that twenty-five-year generational cycles have special importance for the study of genre.

But I also haven't found much evidence for the story of gradual consolidation that I expected to reveal. Although it is clearly true that the publishing institutions governing genre developed gradually, it appears I was wrong to expect that the textual differences between genres would develop in the same gradual way. In the case of detective fiction, for instance, the textual differences that distinguish twentieth-century stories of detection from other genres can be traced back very clearly as far as "The Murders in the Rue Morgue" — and not much farther. Detective fiction did spread gradually, in the sense that Poe and Vidocq were initially isolated figures, without a supporting cast of imitators, let alone genre-specific magazines and book clubs. But textual patterns don't have to develop as gradually as institutions do. Poe's stories already display many of the same features that distinguish twentieth-century crime fiction from other genres.

Predictive modeling

Computers enter this essay largely to address a tangle of problems created by recent genre theory. If we could define genres once and for all by locating a single formal principle that unified them, our critical task would be much simpler. We could say that science fiction is Darko Suvin's "literature of cognitive estrangement" (Suvin, 1972: 372), and be done. Unfortunately, readers rarely agree about the defining characteristic of a genre; different communities may value different things about the same works. Genre theorists increasingly suspect that genres are "family resemblances," constituted by a host of overlapping features

(Kincaid 2003: 413-14). Moreover, genres are historical constructions: the features that matter may change (Rieder 2010: 193).

In short, it increasingly seems that a genre is not a single object we can observe and describe. It may instead be a mutable set of relations between works that are linked in different ways, and resemble each other to different degrees. A problem like this requires a methodology that is cautious about ontological assumptions, and patient with details. Predictive modeling fits the bill. Leo Breiman has emphasized that predictive models depart from familiar statistical methods (and I would add, from traditional critical procedures) by bracketing the quest to identify underlying factors that really cause and explain the phenomenon being studied (Breiman 2001). Where genre is concerned, this means that our goal is no longer to define a genre, but to find a model that can reproduce the judgments made by particular historical observers. For instance, adjectives of size ("huge," "gigantic," but also "tiny") are among the most reliable textual clues that a book will be called science fiction. Few people would *define* science fiction as a meditation on size, but it turns out that works categorized as science fiction (by certain sources) do spend a lot of time talking about the topic. Add clues from a few hundred more words, and you may have a statistical model that can identify other works these same sources called "science fiction," even if the underlying definition of the genre remains difficult to articulate (or never existed).

Hoyt Long and Richard Jean So (2016) have recently used predictive models in a similar way to recognize "latent, nonexplicit traces" of a haiku style in English poetry (266). The point of machine learning in projects like these is not primarily to enlarge the number of books we consider, but to register and compare blurry family resemblances that might be difficult to define verbally without reductiveness. To put it more pointedly: computational methods make contemporary genre theory useful. We can dispense with fixed definitions, and base the study of genre only on the shifting practices of particular historical actors — but still produce models of genre substantive enough to compare and contrast.

Since no causal power is ascribed to variables in a predictive model, the choice of features is not all-important. I'll use words as clues in the discussion that follows, but I don't mean to imply that genre is a linguistic phenomenon. It isn't. Genre is a broadly social phenomenon; words just happen to be convenient predictive clues, allowing us to trace the implicit similarities and dissimilarities between different practices of selection. We could use other features of the text if we preferred. Some researchers have used punctuation marks or character networks to predict genre; when I was trying to locate genres in a collection of 850,000 volumes, I engineered features related to page format (Hettinger et al. 2015; Jockers 2013: 78; Underwood 2014). But our goal in that project was to maximize the sheer predictive accuracy of the model, since we were dragging a net through unknown waters, and wanted simply to catch as much as possible.

This project's goal is different. I am working with labeled examples, not trying to catch unlabeled ones. If all the models described here could be improved by 1%, it would make no difference to the argument. What matter are the *relative* strengths of the boundaries between different groups of texts. So I have made little effort to optimize accuracy; instead I've maximized legibility and consistency. All the models described here use the same feature set, which is created simply by taking the top 10,000 words (by document frequency) in the collection as a whole, across all genres and works without a determinate genre. (We could have used the top 5,000 or 3,000 words; accuracy would vary by about 1%.) For a learning algorithm, I use L2-regularized logistic regression, a well-known algorithm that provides relatively simple estimates of feature importance (Pedregosa 2011; images are produced using Wickham 2009).

This cavalier attitude toward mere accuracy does have a limit. If statistical models couldn't predict genre at all, they obviously wouldn't provide useful evidence. But that's not a problem we will encounter. The models discussed in this article will make predictions that are 70% to 93% accurate, clustering toward the upper end of that range. And although we'll characterize a genre predicted with

only 76% accuracy as a "relatively loose" grouping, compared to one that can be recognized 91% of the time, the truth is that all of these numbers reveal substantial similarities across a group of texts: we're well above ordinary social-scientific thresholds of effect size.[3]

Although I have used the same method for every genre, I cannot guarantee, in advance, that the method will be equally suited to all genres. If a genre was particularly hard to pin down to a vocabulary, it might be hard to classify using a bag-of-words model. Science fiction, for instance, seems likely to pose special problems, since submarines eventually stop counting as futuristic, and are replaced by new wonders of tomorrow. In practice, we won't encounter many problems of that kind. Lexical models have no difficulty finding common formal elements that link thematically diverse works. Generally they report similarities between texts that closely track critical intuition. But they can also diverge from critics' expectations (which are not, after all, in agreement). Positive divergences are easy to interpret: continuities discovered by a lexical model can immediately rule out the thesis that two sites of reception had nothing in common. Our confidence in negative divergences will have to build up more slowly, since we cannot *a priori* rule out similarities that elude the model. These methods will have to recognize the coherence of many different genres before we start to trust that the groupings they see as looser are truly less coherent on a textual level.

If you give a learning algorithm enough variables, it can in effect "memorize" a dataset and make unrealistically accurate predictions about the examples it has already seen. So a model with more than a few variables can only really be tested on held-out examples. The models in this paper are always evaluated by cross-validation on held-out authors (for the rationale see Sculley and Pasanek 2008). In other words, we show the model all the authors (except one) in a

[3] If you convert predictive accuracy into Cohen's *d*, anything above 65% equates to a "large" effect.

set, and then test it on the unseen author's works. The process is repeated until all the authors have been covered and we can calculate accuracy on the whole set.

Genres occupy a space with more than three dimensions

The word "genre" may evoke a mental image of a map that neatly partitions the landscape so that each work is located in one and only one region. But our actual practices of categorizing fiction haven't created that kind of map, nor have I attempted to produce one here. A novel like *The Woman in White* (1859) is assigned by some observers to "the Gothic," and by others to "the sensation novel." In my metadata it bears tags associated with both claims. Other novels aren't associated with any determinate genre: in reality the majority of nineteenth-century works have never been categorized very specifically. So a work of fiction can belong to many genres, or to none at all. Instead of attempting to discover a single partitioning scheme that organizes this whole space, I run a series of separate comparisons, always assembling works tagged with a particular group of genre claims, and comparing that set of works to a contrast set of equal size. Usually the contrast set is selected randomly from a digital library (except inasmuch as it excludes tags in the positive set), and is distributed across time in a way that matches the distribution of the positive set as closely as possible.

If there's no meaningful difference between the two categories being compared, you would expect the model to make predictions that aren't much better than random guessing (50% accurate). And it makes sense to start by running that test as a sanity check. I randomly assigned all the authors used in this article to two "teams" and randomly selected 140 volumes from each "team," then tried to train a model to distinguish the two groups. The results of forty trials are plotted (in gray) in figure 1. As you can see, average accuracy is a little lower (45%) than it would be if we had just guessed randomly. When a classification algorithm tries to find differences between two randomly-selected groups, it is still able to discover a faint pattern, but an accidental pattern will have no meaningful relation

to held-out examples, and a useless rule of that kind can easily give predictions a bias that is worse than a guess. So there is very little danger that this algorithm will seem to discern a difference between two groups where none exists.

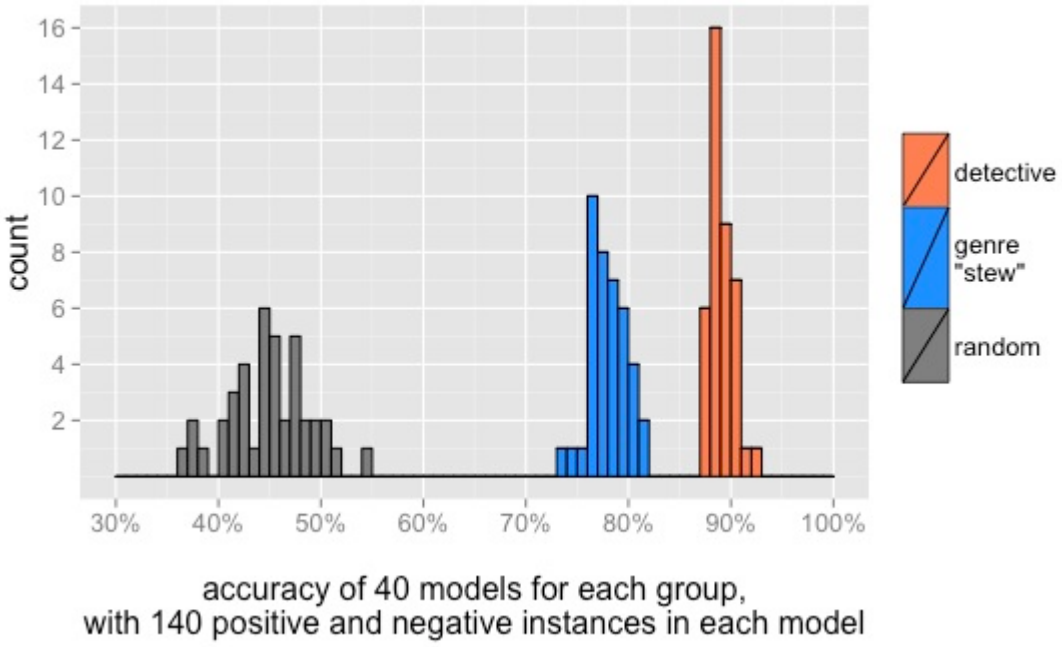

Figure 1. Histogram plotting the accuracy of 40 models for three different putative "genres." For each model, 140 positive instances were selected randomly from a longer list. The list of "detective fiction" was constructed using methods to be explained in the next section.

But there is another kind of baseline test we should run. What happens if we mix all the works tagged with *any* genre we're studying into a single ghastly stew and compare that superset to all the randomly-selected works that weren't associated with any genre tag? As you can see in Figure 1, the model is often able to recognize volumes that come from our "genre stew," even though this combination of Gothic, Newgate, sensation, detective, and science fiction probably doesn't constitute anything we would ordinarily call a coherent genre. The model's

predictions are correct, on average, 78% of the time.[4] Glancing at a few of the words that predict membership in this superset of genres, it's not hard to see what's happening. The individual genres involved here are not entirely dissimilar. They share sensational subject-matter, and a number of props or plot devices that are more likely to occur in any of them than in a randomly-selected work. "Murder," "ghastly," "lock," "key," "theory," and "laboratory" are near the top of the list of predictive words, for instance. The words that typify the random contrast set are harder to characterize, but (by comparison at least) evoke aspects of ordinary domestic life ("married," "blame," "mornings," "proud," "friends," "afternoons"). Perhaps this picture would change if we were studying genres like the Bildungsroman. But in the dataset I have assembled for this article, there are broad differences between genre fiction as such and a randomly-selected, relatively quotidian background.

If we wanted to understand this difference in depth, we would need to do more than glance at the top and bottom of a list of ten thousand features. A semantic scaffolding could no doubt be built to support and particularize my casual inferences about "sensational" and "quotidian" subject-matter. But that would take up space and time, and it is not the primary point of this essay to offer new descriptions of the content of every genre it discusses. So, while full lists of features are available in an online code and data supplement (Underwood 2016), my descriptions will remain brief, only mentioning a few predictive words from each model to convey a general flavor of the contrast involved.

Instead of redefining genres, this essay is fundamentally making an argument about their varying lifespans and degrees of textual coherence. For that purpose, what matters is less how we characterize diction, and more how we interpret degrees of similarity between groups of texts. It is especially important to understand that genres can occupy a space of similarity with more than three

[4] Since we will always be comparing two evenly-sized categories in this paper, I won't bother distinguishing precision from recall.

dimensions. Things that are close to each other along one axis can still be far apart in other ways. For instance, we have seen that detective fiction, science fiction, and various other genres, mixed together, can be distinguished from randomly-selected works 78% of the time. It might be tempting to infer that detective fiction and science fiction are "basically continuous" or "78% similar"; we could start to build a narrative that made Sherlock Holmes' training in chemistry a crucial, overlooked connection between them. But in fact, the differences *between* these genres are even stronger than their collective difference from a randomly-selected background. If you train a model to distinguish detective fiction from science fiction, or from the Gothic, it will be right (in both cases) more than 93% of the time.

Predictive models are rather like human beings: they can always find some ways that that two sets of works are similar, and other ways that they differ. If we want to know whether detective fiction and science fiction can usefully be lumped together, no single two-sided comparison will answer the question. Instead we might look at the relative strength of multiple comparisons. We see in figure 1, for instance, that detective fiction on its own is significantly easier to distinguish from a random contrast set than our "genre stew"; this might already suggest that it's a more tightly-knit category. Alternatively, we might ask a three-sided question that allows us to situate two genres in the same frame of reference. Do detective stories, for instance, differ from other works of fiction *in the same way* that science fiction differs? Predictive models are good at extrapolating from one set of evidence to another. So you can train a model on the contrast between detective fiction and a randomly-selected background, and then ask the same model to distinguish works of science fiction from the same background. As we might expect, the model fails utterly: it's right less than half of the time. Although these two genres have a few things in common (theories and laboratories, for instance), *most* of the features that distinguish them from the background are different. When I need to decide whether two models of genre are similar, in the pages that follow, this is the test I'll place most trust in. It certainly tells us that our ghastly "genre stew" can be

separated out into detective fiction and science fiction. The next question is whether "detective fiction" itself similarly breaks up into subgenres, or groups of works assembled at particular sites of reception, that differ more than they resemble each other.

Detective fiction

In assembling genre lists for the second half of the twentieth century, I have used mostly Library of Congress genre/form headings. These tags were applied to volumes by individual librarians, and reflect tacit assumptions about genre held by many different people. In the case of detective fiction, I've lumped together several different headings, including "Mystery fiction," "Detective and mystery fiction," and the subject heading "Detectives" (when applied to works that are mostly fiction). As we go back before 1940, these tags become very sparse, because we're looking at works that were originally cataloged before the Library of Congress system assumed its present form. Only a few of these works have been recataloged in the modern verbose way. So for the earlier period we mostly have to rely on bibliographies and critical studies.

There are many enormous bibliographies of detective fiction; the challenge is to find one small enough to transcribe. For pre-war detective fiction, I have relied mostly on the catalog of an exhibition organized at Indiana University Library in 1973, covering "The First Hundred Years of Detective Fiction, 1841-1941." This exhibition lists collections of short stories (and a few individual stories) along with novels. It also lists works in translation. I have been similarly inclusive throughout this essay. A writer like Jules Verne did an enormous amount to shape genre beyond France, so we would lose a lot by excluding translations. I doubt, moreover, that there's anything untranslatable about the patterns at issue here: Verne will turn out to be an extremely typical figure within science fiction, even in translation.

It is of course possible that a single exhibition catalog of detective fiction (limited to volumes before 1941) will create a picture of the genre that diverges substantially from assorted volumes 1829-1989 cataloged by many different hands. But that's exactly the kind of question statistical modeling allows us to test. Modeling just the 88 volumes from the Indiana exhibition (that I was able to obtain digitally), we have a rather high level of accuracy, 90.9%. The 177 volumes that have Library of Congress genre tags are more of a mixed bag, and can only be recognized 87.3% of the time. If we combine both sets, we have 249 volumes (since 16 were in both groups) that can be recognized with 91.0% accuracy. So mixing groups selected in different ways doesn't reduce accuracy; it's a compromise that "levels upward."

But as I've mentioned before, algorithmic models can be very good at finding common elements in a group of works. The real test of similarity between two categories (A and B) is to ask a model trained on a contrast between A and C to also distinguish B from C. For instance, when we ask a model trained on the Library of Congress detective fiction to distinguish the Indiana exhibition from a similar random background, it is still 89.3% accurate. That's the real confirmation that we're looking at largely congruent definitions of detective fiction.

The probabilistic nature of the model we're using makes it easy to see which examples of detective fiction are particularly typical or particularly hard to classify. We can spread volumes out along a y-axis that characterizes the model's degree of confidence that they belong to the "detective fiction" set. In Figure 2 I've done that with all 249 volumes that were either tagged by individual librarians or included in the Indiana exhibition.

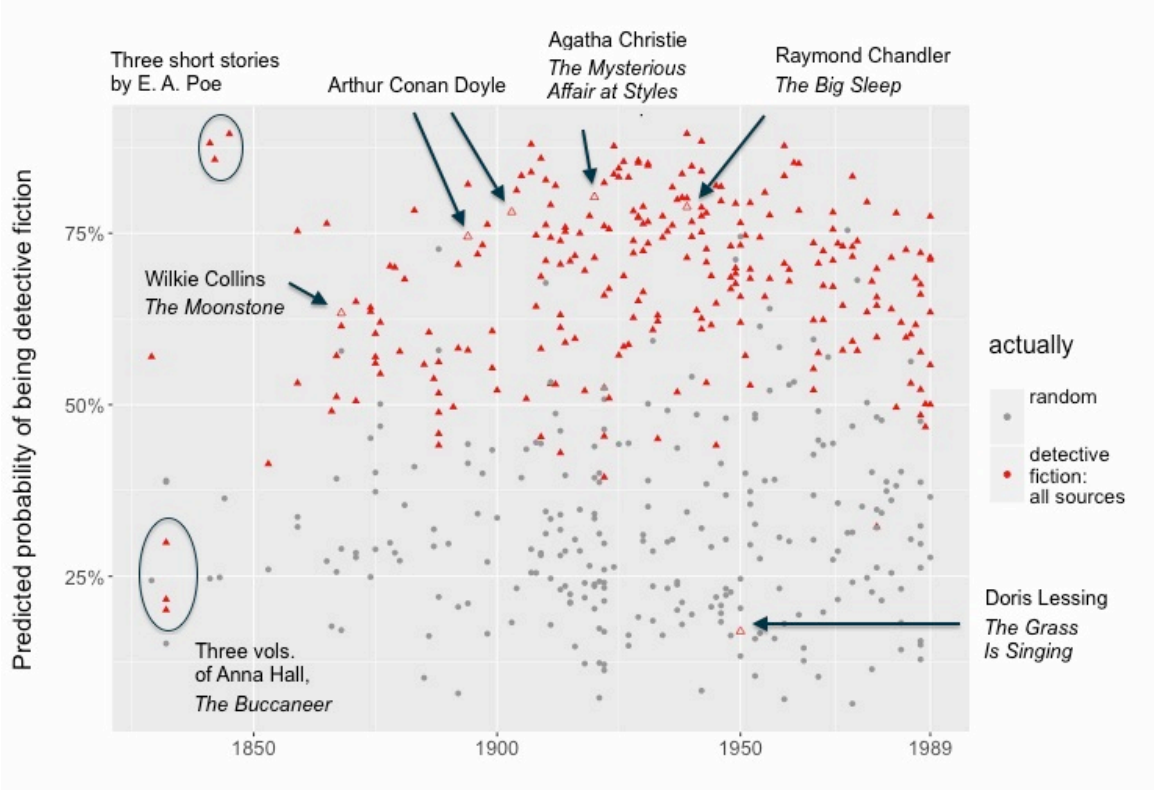

Figure 2. Predicted probabilities of coming from the "detective fiction" set; 91.0% accuracy overall.

One very striking detail is the position occupied by Edgar Allan Poe's three stories of detection from the early 1840s. They seem to be exemplary models of the genre, not just in their own period, but according to standards that organize the whole timeline from 1829 to 1989. Although Poe's durable status as a template is consistent with one influential genealogy of the detective story (Rachman 2010), it is not something all critics have agreed about. Moretti, for instance, remarks that detective fiction achieved its modern form only around 1890 (2005: 31), and Ascari outright denies that Poe is still relevant to crime fiction (2014: 15). Nor is this continuity something I actually expected to see in a statistical model. In fact, I expected that the boundaries of detective fiction would tend to get blurry as we proceeded back before Conan Doyle, into a period where stories of detection were often fused with other genres, like the sensation novel. Perhaps we see a bit more

blurriness in the 1860s and 1870s than in the middle of the twentieth century. (And there are some flat-out errors in the 1830s: catalogers who tried to stretch "detective fiction" to cover a novel from 1832 end up breaking the concept.) But the early works critics tend to identify as prototypes ("Murders in the Rue Morgue," *The Moonstone*) remain exemplary in this model. This evidence doesn't necessarily establish an "origin," or prove that particular writers defined the genre. It might prove only that the late-twentieth-century critics who identified prototypes of detective fiction did a good job of extrapolating backward from practice in their own era.

But this evidence does show that the continuity of detective fiction is more than a sequence of genealogical links connecting disparate forms. For instance, suppose we lump the Indiana exhibition and Library of Congress tags into a single group of 249 texts, but divide the group chronologically at the year 1930. How much does the definition of detective fiction change between the two halves? If we model the 130 volumes after 1930, we get 88.5% accuracy. But if we train a model on the 119 volumes up to 1930, and use that model to make predictions about works after 1930, the model will still be 86.9% accurate. The verbal differences that mark detective fiction up to 1930 largely continue to characterize it afterward. This is not to say that the genre stopped changing. The main vogue of the hard-boiled detective, for instance, is still to come in 1930; that's certainly an important change. But the nature of the *difference* between detective fiction and the rest of the literary field didn't dramatically alter. The boundaries of the genre are stable. (It may be worth noting that this remains true even though the random contrast set before 1930 comes mostly from HathiTrust, and after 1930 mostly from the Chicago Text Lab. These collections were selected differently, but the differences are not large enough to interfere with the genre signal.)

Our model provides evidence of continuity strong enough to pose real problems for a prevailing strain of nominalism in genre theory. The (valid) premise that genres needn't be unified by a clear definition or pre-existing essence is often

taken a step or two further, to suggest that genres are unified only by a genealogical thread — that past and future are linked only as a continuous process of negotiation among "communities of practice" (Rieder 2010: 203). Maurizio Ascari criticizes Moretti's "positivistic" approach to detective fiction by reminding us that "in the course of the twentieth century, detective fiction deeply changed"; indeed, "all literary genres change unceasingly" (Ascari 2014: 7, 15). Mark Bould and Sherryl Vint (2009) argue that "genres are never, as frequently perceived, objects which already exist in the world and which are subsequently studied by genre critics, but fluid and tenuous constructions made by the interaction of various claims and practices" — or putting it even more boldly, "there is no such *thing* as science fiction" (48, 43). When I began this study, I might have embraced some of these claims; at any rate, skepticism about the stability of genre seemed preferable to endless definitional argument. But predictive models make a middle path possible. We can start cautiously, with contingent boundaries drawn by specific historical actors, and then ask, empirically, how far their implicit selection criteria agree or diverge. In the case of detective fiction, lists of texts organized by different hands, at different times, are extremely compatible. Moreover, a model of the genre's past does an excellent job of predicting its future.

But what exactly *is* the definition of "detective fiction" operative here? It is not, for the most part, a shocking one. "Police," "murder," "investigation," and "crime" define the thematic premise of the genre. "Suspicion," "evidence," "prove," "theory," "coincidence" (and, to give a subtler example, "whoever") foreground the mechanics of doubt and demonstration that drive the plot. If we look a little deeper into the model, there are less obvious details. For instance, architecture and domestic furnishing also provide clues: "door," "room," "window," "desk" are all highly predictive words. At the opposite end of the scale, words that describe childhood and education ("born," "grew," "taught," "children," "teacher") strongly predict that a volume is not detective fiction.

Perhaps the genre's focus on a particular mysterious incident (or its tendency to take short-story form) encourage a contraction of biographical horizons.

In any event, detective fiction turns out to be textually coherent across a period of 160 years (1829-1989). But we haven't really tested a model until we find out where it breaks. For instance, suppose instead of foregrounding the figure of the detective, we foregrounded the criminal milieu? The boundary between "detective fiction" and "crime fiction" can be blurry; it is troubled by the novels of Patricia Highsmith and by Arthur J. Raffles, gentleman thief. What if we added a couple of novels that are tagged as "crime fiction" but not detective fiction in the twentieth century, and also a group of Newgate novels from 1820-40? One classic study of detective fiction begins with *Oliver Twist,* after all (Miller, 1988). Sensation novels have also been identified as precursors of detective fiction (Pittard, 2003); *The Moonstone* was already included in the Indiana exhibition, but we might try adding other sensation novels to see whether they too would fit in this category.

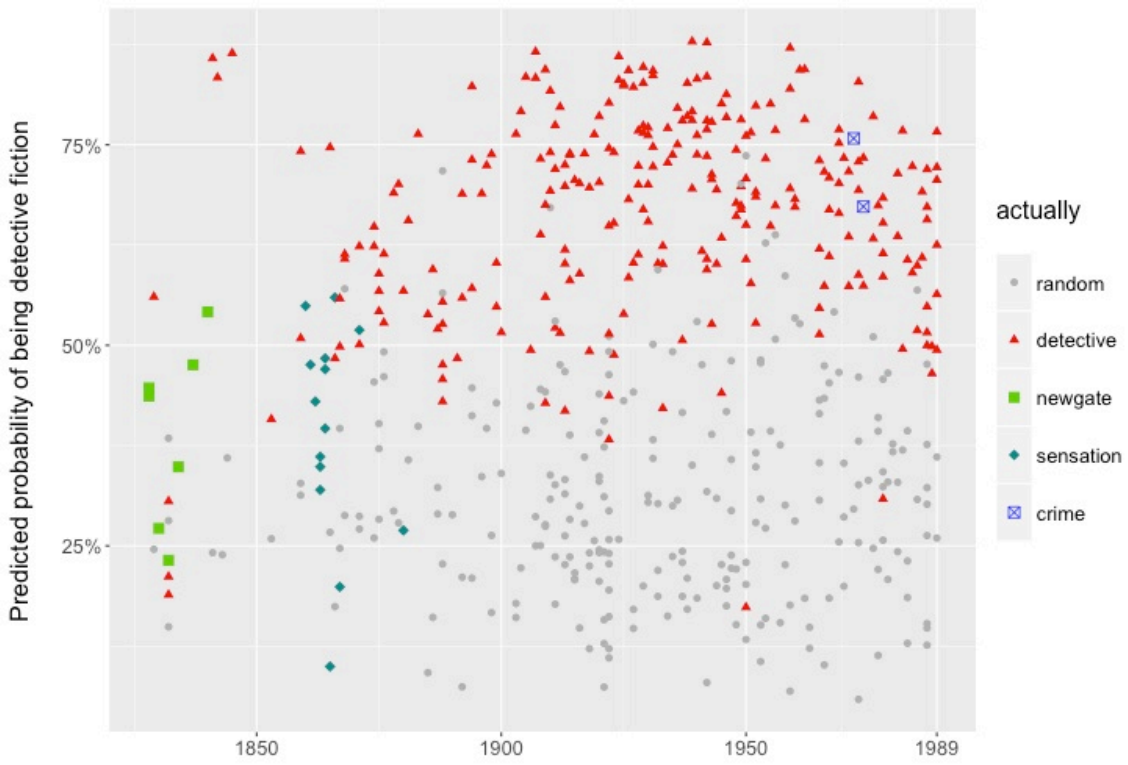

Figure 3. A model trained only on the detective and random examples makes predictions about three other categories.

There are places where our model can easily include other categories and places where it refuses to stretch. In figure 3, I've trained a model on the same group of 249 detective novels used in figure 2, but have also allowed it to make predictions about other sets of works not included in the training set. Recent novels that were tagged as "crime fiction" (for instance Patricia Highsmith, *A Dog's Ransom*) turn out to be very compatible with our existing model of detective fiction. But nineteenth-century sensation novels and Newgate novels won't fit into the same textual box. This doesn't prove that it was wrong for D. A. Miller to discuss Newgate and detective novels together in *The Novel and the Police*. There is no law, after all, declaring that literary-historical concepts have to be recognizable at the level of diction. If we want to define a genre called crime fiction that includes the Newgate novel, we can do it, and the concept may well be illuminating. But we will be talking about a genre of a slightly different kind — one that lacks the level

of linguistic homogeneity connecting E. A. Poe to Agatha Christie and Patricia Highsmith. The point of this inquiry, in other words, is not to decide what can or cannot be called a "genre," but to help us differentiate the various kinds of patterns literary historians have used the word to designate.

The Gothic

The history of "Gothic fiction" creates, appropriately, a mystery about ancestral figures that haunt their descendants only as ambiguous traces. Critics seem fairly confident that the Gothic novel was a coherent phenomenon in Britain from 1760 to perhaps 1830. But as we move further into the nineteenth century, it becomes less and less clear whether the Gothic remains a continuous tradition. I have already mentioned that Franco Moretti divides the nineteenth-century Gothic into two genres at opposite ends of the century. In twentieth-century America, "Southern Gothic" is often treated as a distinct literary phenomenon. There's also a strong argument to be made for a specifically female Gothic tradition that might run back through DuMaurier's *Rebecca* to Brontë's *Jane Eyre* (Fleenor 1983). On the other hand, there are critical traditions that insist on the continuity of all these things, and that indeed stretch the Gothic to encompass the contemporary publishing category of "horror" (Edmundson, 1999). You can find anthologies of Gothic fiction for sale that span the whole distance from *Otranto* to Anne Rice, so as a practical matter of reception, there must be *some* sort of continuity out there, whether we want to call it a genre, a mode, a fandom, or a loose set of themes.

Since we have good reasons to wonder whether "the Gothic" writ large is a strongly unified tradition, it was particularly important in this case to compare different sources of testimony. Before 1840 I relied heavily on the Stanford

Literary Lab's list of Gothic fiction;[5] after 1940 I relied increasingly on the Library of Congress genre tags associated with "horror" or the "ghost story." But I also collected a set of works mentioned in *The Gothic,* a Blackwell guide edited by David Punter and Glennis Byron (2004), which tries to trace a Gothic tradition all the way from Horace Walpole to Brett Easton Ellis, linked through a surprising range of intermediary figures that includes Charlotte Brontë, Henry James, and H. P. Lovecraft.

None of these lists display the kind of coherence we found in detective fiction. The sample that could be predicted most accurately was the smallest: the 21 works (1791-1834) identified as Gothic by the Stanford Literary Lab could be recognized 81.0% of the time. The hardest sample to model was the superset that combines them all: 165 volumes that can only be recognized with 77.0% accuracy. The numeric contrast here between 81% and 77% is a little more dramatic than it sounds, because these aren't apples-to-apples comparisons. Accuracy would ordinarily increase with the size of the set being modeled. To convey a sense of that increase, I've plotted curves that indicate the mean accuracy for other genres at various sample sizes.

[5] The metadata I use for the "Stanford Gothic" were developed at the Stanford Literary Lab; many hands may have been involved, including certainly those of Ryan Heuser and Matthew L. Jockers.

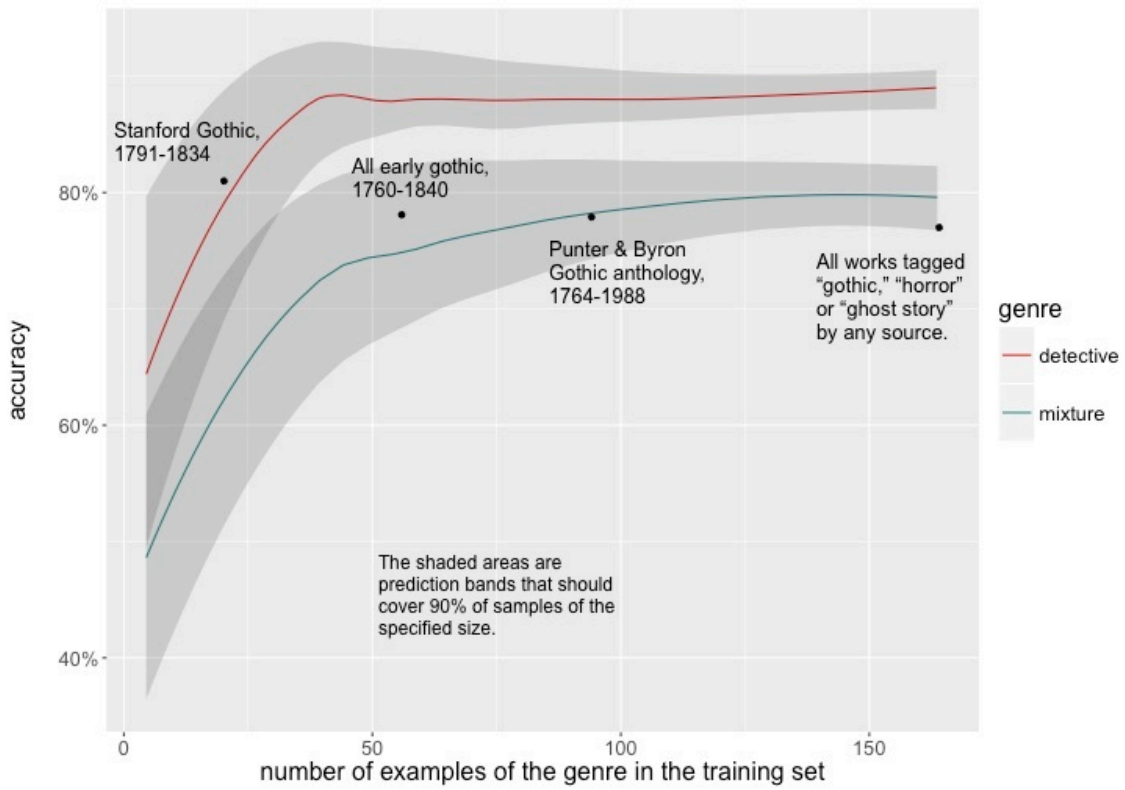

Figure 4. Various samples of Gothic fiction, plotted relative to curves that indicate the typical range of accuracies for other genres at different sample sizes.

The Stanford subset of Gothic novels is textually as coherent as a similarly-sized sample of detective fiction. But century-spanning samples of Gothic fiction perform remarkably poorly for their size; they're no easier to model than a mixture of all genres included in the project. This suggests that the Romantic-era Gothic novel has very little language in common with twentieth-century traditions of horror or supernatural fiction. We can confirm this by training a model on one of these periods and applying it to the other; it does little better than random guessing.

I don't expect this to surprise many readers. Very few critics claim that the Gothic, writ large, is a genre as tightly-knit as the sensation novel or detective fiction. Even in the process of constructing a two-century anthology, Punter and Byron (2004) acknowledge the possibility that "there are very few actual literary texts which are 'Gothic'; that the Gothic is more do to with particular moments,

tropes, repeated motifs that are found scattered … through modern western literary tradition" (xviii). The difficulty of modeling Gothic fiction on a two-century scale is partly a confirmation of this suspicion, but also perhaps a sanity check for our method. (A method that didn't recognize hard cases would be hard to trust about others.) It may also remind us that there are continuities a model based on diction fails to register. Critics seem to agree that the spectrum of things called Gothic constitutes, if not quite a genre, at least a mode or a loose thematic similarity. Yet our model doesn't see that spectrum as unified more strongly than works selected at random from an assortment of popular genres.

One obvious explanation might be that the concept of Gothic simply covers too much space on the timeline; language just changes too much in two centuries for genre to be modeled linguistically across that distance. But we have other examples where this doesn't seem to pose a problem — for instance, figure 4 reminds us that detective fiction holds together quite well from 1829 to 1989. Indeed, as we see in figure 5, there's not much evidence that chronologically focused genres are generally more coherent, linguistically, than our 160-year sample of detective fiction.

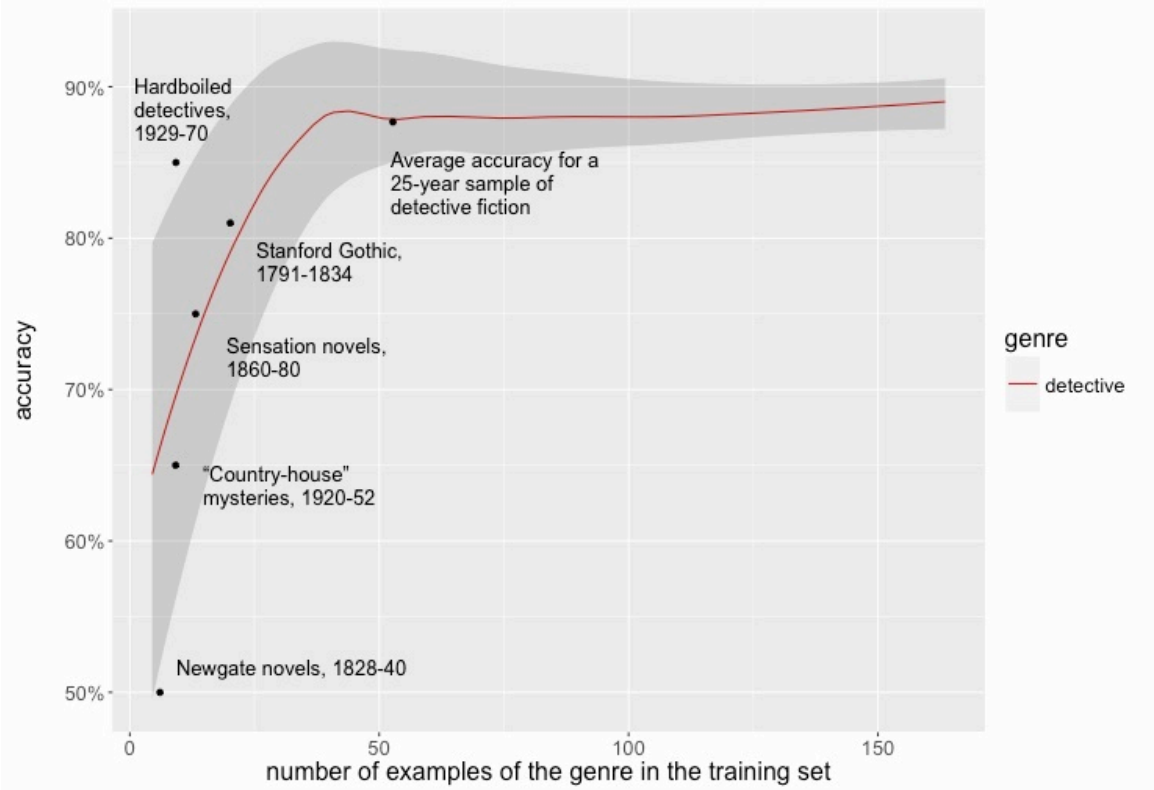

Figure 5. Several genres of roughly generational size, plotted relative to a curve that indicates the range of accuracy for a random sample of detective fiction drawn from 1829-1989. The shaded ribbon is a predictive band that covers 90% of models for a random sample of detective fiction.

If you downsample detective fiction to get a training set the same size as a sample of Gothic, Newgate, or sensation fiction, then detective fiction appears just as coherent as those chronologically-focused genres, even though its volumes are drawn from a period lasting more than 150 years. This is the decisive evidence against Franco Moretti's conjecture that genres have generational lifespans. Genres that survive much longer than a generation seem to be united by textual similarities just as strong as those uniting shorter-lived ones. Since it's hard to prove that these models are capturing all possible similarities between texts, we might express this cautiously: if there is any generational rhythm in the history of genre, this method has not detected it — although it seems able to detect all of the patterns scholars call genres.

Is it possible to make detective fiction even more coherent by focusing on a narrower generational span? Not really. If you break the long arc of detective fiction into 25-year spans of time, and model them separately, you get an average accuracy that is very comparable to the accuracy of a similarly-sized sample from the whole timeline. You can, however, get slightly higher accuracy from at least one critically recognized subgenre: the hardboiled detective novel. Even a small sample of ten hardboiled novels from 1929-1970 can be picked out of a lineup with 85% accuracy. Perhaps the stylistic homogeneity of these novels has something to do with their social homogeneity. Dorothy Hughes is the only woman in our mid-century sample of hardboiled writers. In fact, the distinction between the hardboiled and country-house traditions was not a generational divide at all; it was organized more by gender and by the Atlantic Ocean. But it is also not an insuperable divide. A model trained on the country-house tradition can spot hardboiled detectives just as accurately (85%) as a model trained on the hardboiled examples themselves. So there is little reason to conclude that these subgenres became independent of the larger concept "detective fiction."

Science fiction

I've argued that genre concepts that persist for more than a century can be just as coherent, linguistically, as those that persist for a few decades. But so far detective fiction is my only example, and there are reasons to suspect that the detective/mystery/crime genre might rely on an unusually stable set of premises. There's always a crime; there's always a detective; there's always an investigation. Science fiction would appear to pose a more challenging problem, because the premises of the genre are inherently mutable. The Vernian prototypes of the genre often describe conveyances like balloons and submarines that are no longer science-fictional. Recent examples of the genre depend on technologies of a very different sort. It's not immediately obvious that Gibson's *Neuromancer*, Wells' *Time Machine*, and Mary Shelley's *Frankenstein* would share much common vocabulary.

Many skeptical theories of genre have taken shape specifically around the mutability of science fiction — as the title "There Is No Such Thing As Science Fiction" reminds us (Vint and Bould 2009; see also Kincaid 2003, Rieder 2010). In short, it is not intuitively clear whether we should expect science fiction to hold together over long timelines, like detective fiction, or fall apart like the Gothic.

Since there are different stories about the early history of science fiction, I drew on several different sources for that period. *The Anatomy of Wonder* is a well-known bibliography with chapters on the early history of science fiction contributed by Brian Stableford, a writer of science fiction himself. Stableford's history of the genre strongly emphasizes H. G. Wells and the future-war tradition, but is somewhat more reticent about other predecessor figures, like Mary Shelley and Jane Loudon. (Like many historians of science fiction, Stableford tends to define the genre through its scientific content, and he can be skeptical about works where that content seems lacking.) To get a fuller representation of women in the genre I relied on a bibliography of women in early science fiction constructed by Mary Mark Ockerbloom at the University of Pennsylvania library. In spite of their different conceptual emphases, these sources construct lists of texts that can be modeled, linguistically, in very similar ways. A model of either one, combined with twentieth-century texts labeled by librarians, can predict the other with 90% or better accuracy.

When all these bibliographic sources are folded together, we have a list of 196 volumes stretching from 1771 to 1989 that can be modeled with 88.3% accuracy. The boundaries of the genre are a little less clear than detective fiction, but it certainly has a coherence more akin to that genre than to the Gothic. (Accuracy remains about three times closer to detective fiction even if we downsample all the genres to have the same number of volumes.)

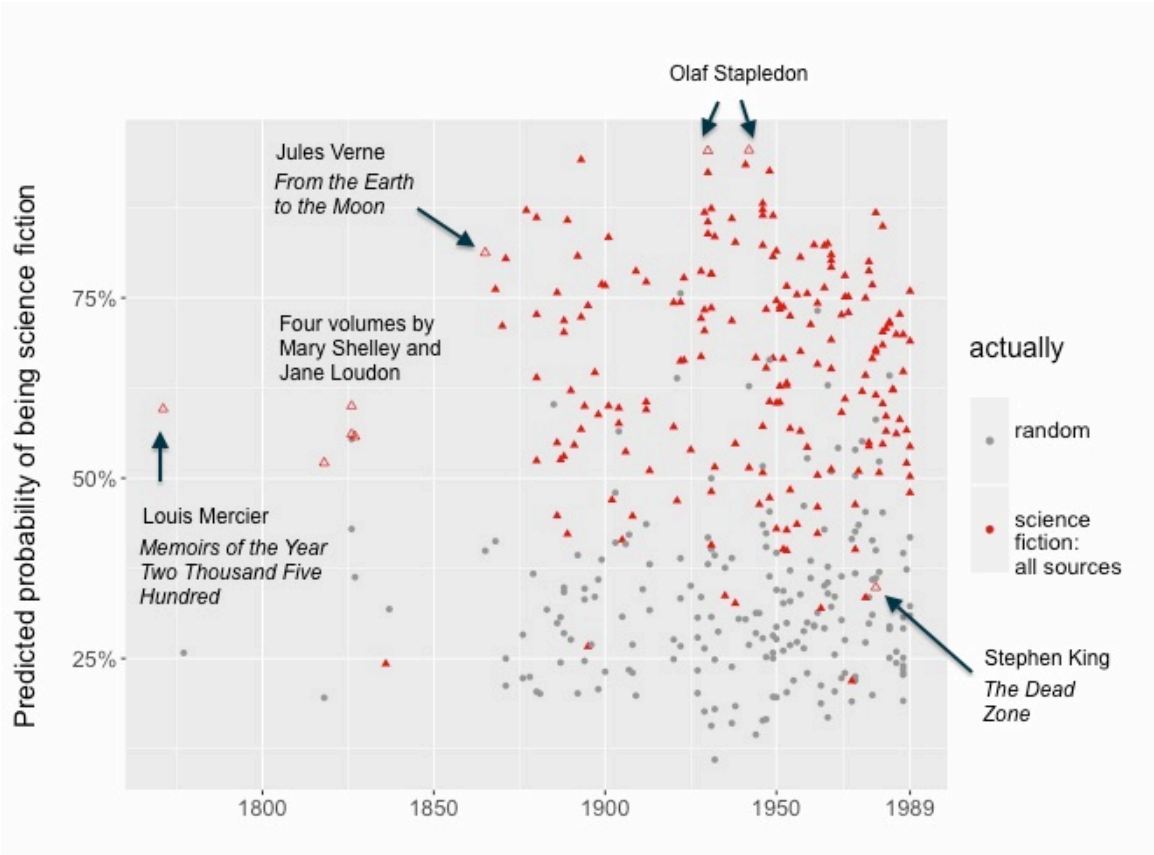

Figure 6. Science fiction 1771-1989, classified with 88.3% accuracy.

How is it possible to model such a long and protean history just by counting words? *Frankenstein* is a Romantic-era text that doesn't even invoke "science" terribly often. But it turns out that there are verbal traces that do persist across 170 years. Invocations of scale ("vast," "far," "larger") are very characteristic of science fiction, as are large numbers ("thousands"). Self-conscious references to the "earth" and to things that are "human" tend to accompany "creatures" from which humanity may be distinguished, and the pronoun "its" is common, since we often confront actors who lack an easily-recognized human gender. This is not by any means an exhaustive description of the genre — just a taste of the model.

One visually salient thing about figure 6 is a slight downward slope in the red triangles after 1950. It's far from clear that this is a statistically significant trend, but it does interestingly echo a similar trend in detective fiction (fig. 3). In

both cases, it appears that the postwar volumes gradually lose the strongly-marked generic distinctness that typified the 1930s and 40s. There are reasons to be cautious here; it's a subtle trend, and it could be shaped by vicissitudes of selection, since mid-century volumes are sometimes more likely to be available digitally if they possess a genre-specific fan audience. It's also conceivable that works in the middle of a timeline might tend to fit a model better than works on either edge (although that pattern hasn't been evident in previous research using this method). But, if the trend is real, it might echo Gary Wolfe's hypothesis that the boundaries of fantastic genres have recently become unstable. "Fantasy is evaporating … growing more diffuse, leaching out into the air around it, imparting a strange smell to the literary atmosphere" (Wolfe 2011: viii). There are however slight differences from Wolfe's thesis, and from related claims about postmodern "magic realism." The shift visible in these models seems to have made genre fiction more like the rest of the literary field. The converse trend — a playful borrowing of genre tropes by mainstream literary authors — isn't particularly visible yet. Perhaps it would become visible after 1990.

Another thing we might expect that doesn't appear in this model is the gradual consolidation of genre conventions that science fiction scholars spend so much time tracing. Historians of this genre are rarely as willing to give Verne and Shelley quite as much credit as historians of detective fiction give Poe. The narrative premise of much historiography is that science fiction was an inchoate phenomenon (scattered across utopias, planetary romances, etc) until given a new shape and direction by particular pulp magazines and anthologies between 1925 and 1950. Hugo Gernsback's *Amazing Stories* (1926) often plays a central role. Wolfe says, for instance, that "science fiction, despite its healthy legacy throughout the nineteenth century, was essentially a *designed* genre after 1926" (34). Even after that point, "the science fiction novel persistently failed to cohere as a genre in the manner of mysteries and Westerns" until *The Pocket Book of Science Fiction* emerged in 1943 (21). None of these crucial moments of consolidation are visible in the model.

Where language is concerned, the half-century from Verne through Gernsback (1875-1925) appears just as coherent and as distinct from other forms of fiction as the period after 1926. It's possible, of course, that the model is wrong. Maybe the mere linguistic distinctiveness of science fiction is not as important as other forms of consolidation. But it also seems possible that the historiography of science fiction has been unduly impressed by Gernsback's coinage of the term "science fiction," or by the romance of the pulp era, or by the symbolic centrality of certain technologies (like spaceflight), and has tended to undersell the genre's fundamental coherence with earlier traditions of scientific romance.

What have we learned?

The evidence gathered in this article challenges three existing theories of genre. Franco Moretti's conjecture that genre is a generational cycle is probably the least important of these targets: Moretti offered it as a reluctant speculation, and it has only been adopted by a few other critics (Roberts 2011). The premise that genre boundaries gradually "consolidate" in the early twentieth century is a more serious matter. The notion that the pulps gave form to protean traditions that had previously "failed to cohere" is very influential in science fiction criticism (Wolfe 2011: 21). I don't think I have refuted this notion yet, but linguistic models provide at any rate a striking lack of evidence for it: the distinctive language of science fiction seems to take form before the institutions that are supposed to have consolidated it. The third theory of genre I have questioned is the recently popular notion that histories of genre are merely a genealogical thread linking disparate cultural forms. Predictive models can directly challenge this claim. If a model trained on detective fiction before 1930 can also recognize detective (and crime) fiction after that date, then then the differences separating the genre from the rest of the literary field must have remained relatively stable.

But I should emphasize again that stable generic *boundaries* are not the same thing as a stable *definition* of the content inside the boundary. Although the

particular words associated with genres are often fascinating, I have treated those details deliberately casually here, to avoid implying definitional claims. In fact, this essay relies on predictive models exactly because they can bracket questions of definition, and start instead from the wary, nominalistic premises that underpin contemporary skepticism about genre. For instance, predictive models easily embrace the notion that genres involve "family resemblances" composed of many features rather than a single defining characteristic (Kincaid, 2003). Predictive models are also quite compatible with the assumption that genres are constituted by competing, subjective acts of "labeling," rather than deep formal structures waiting to be revealed (Rieder 2010: 192-93). This article relies on computational methods because they allow us to build on this sort of plural and perspectival foundation. Otherwise it would have been difficult — for me at any rate — to characterize and compare the strength of complex family resemblances traced by many different observers.

But even an inquiry that begins from perspectival premises can end up revealing that competing acts of labeling were actually, in some cases, implicitly compatible. And even an inquiry founded on the premise of historical mutability can turn out to show that the boundaries of some genres remained stable for a century and a half. I think that is what we have seen with detective fiction and science fiction (although perhaps both genres do begin to "evaporate" by the end of the twentieth century).

We cannot expect to see the same stability in every case. The cluster of phenomena called Gothic, for instance, are more reluctant to coalesce. Many nineteenth-century genres (like Newgate and sensation fiction) do seem to be as short-lived as Moretti claimed. Even science fiction is slightly more protean than detective fiction. If there is a single central thesis to be drawn from this paper, it is that the things we call "genres" may be entities of different kinds, with different life cycles and degrees of textual coherence. Literary scholars have groped toward some acknowledgment of this by distinguishing "genres," for instance, from

"modes." But that binary distinction seems insufficient to describe the historical continuum explored here. Although this article rejects Franco Moretti's conjecture that genres have a generational rhythm, I think it vindicates his broader contention that quantitative methods can give literary scholars descriptive resources that are more flexible and more responsive to the complexity of our material.

References

Ascari, Maurizio. 2014. "The Dangers of Distant Reading: Reassessing Moretti's Approach to Literary Genres." *Genre* 47, no. 1: 1-19.

Bould, Mark, and Sherryl Vint. 2009 "There is No Such Thing As Science Fiction." In *Reading Science Fiction,* edited by James Gunn, Marleen Barr, and Matthew Candelaria, 43-51. New York: Palgrave.

Bourdieu, Pierre. 1984. *Distinction: A Social Critique of the Judgment of Taste*. Cambridge: Harvard UP.

Breiman, Leo. 2001. "Statistical Modeling: The Two Cultures." *Statistical Science* 16, no. 3: 199-231.

Charles, John, Joanna Morrison, and Candace Clark. 2002. *The Mystery Readers' Advisory: The Librarian's Clues to Murder and Mayhem*. Chicago: ALA.

Devitt, Amy. 2004. *Writing Genres.* Carbondale: SIU Press, 2004.

DiMaggio, Paul. 1987. "Classification in Art." *American Sociological Review* 52: 440-455.

Edmundson, Mark. 1999. *Nightmare on Main Street: Angels, Sadomasochism, and the Culture of Gothic.* Cambridge, MA: Harvard University Press.

Fleenor, Juliann E. ed. 1983. *The Female Gothic* Montréal: Eden.

Hettinger, Lena, Martin Becker, Isabella Reger, Fotis Jannidis, and Andreas Hotho. 2015. "Genre Classification on German Novels." 12th International Workshop on Text-Based Information Retrieval. http://www.uni-weimar.de/medien/webis/events/tir-15/tir15-papers-final/Hettinger2015-tir-paper.pdf

Hughes, Winifred, 2002. "The Sensation Novel." In *A Companion to the Victorian Novel,* edited by Patrick Brantlinger and William B. Thesing, 260-78. Oxford: Blackwell.

James, P. D. 2011. *Talking About Detective Fiction*. New York: Vintage.

Jockers, Matthew L. 2013. *Macroanalysis: Digital Methods and Literary History*. Urbana: University of Illinois Press.

Kincaid, Harold. 2012. "Doing Philosophy of Social Science." In *The Oxford Handbook of Philosophy of Social Science,* edited by Harold Kincaid. Oxford: Oxford University Press. http://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780195392753.001.0001/oxfordhb-9780195392753-e-1

Kincaid, Paul. 2003. "On the Origins of Genre." *Extrapolation* 44, no. 4: 409-419.

Knox, Ronald. 1929. "Introduction," *The Best Detective Stories of 1928-29*, ed. Ronald Knox. Reprinted in Haycraft, Howard, *Murder for Pleasure: The Life and Times of the Detective Story,* Revised edition, New York: Biblio and Tannen, 1976.

Long, Hoyt, and Richard Jean So. 2016. "Literary Pattern Recognition: Modernism between Close Reading and Machine Learning." *Critical Inquiry* 42, no. 2: 235-67.

Miller, Carolyn. 1984. "Genre as Social Action." *Quarterly Journal of Speech* 70: 151-167.

Miller, D. A. 1988. *The Novel and the Police* Berkeley: University of California Press.

Moretti, Franco. 2005. *Graphs, Maps, Trees*. London: Verso.

O'Brien, Geoffrey. 1981. *Hardboiled America.* New York: Van Nostrand Reinhold.

Ockerbloom, Mary Mark. 2015. "Pre-1950 Utopias and Science Fiction by Women." http://digital.library.upenn.edu/women/_collections/utopias/utopias.html Accessed 2015.

Pedregosa et al. 2011. "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research* 12: 2825-2830.

Pittard, Christopher. 2010. "From Sensation to the *Strand.*" In *A Companion to Crime Fiction*, edited by Charles J. Rzepka and Lee Horsley, 105-116. Oxford: Wiley-Blackwell.

Punter, David, and Glennis Byron. 2004. *The Gothic.* Malden: Blackwell.

Rachman, Stephen. "Poe and the Origins of Detective Fiction." In *The Cambridge Companion to American Crime Fiction*, edited by Catherine Ross Nickerson, 17-28. New York: Cambridge University Press, 2010.

Roberts, Adam. "A Brief Note on Moretti and Science Fiction." In *Reading Graphs, Maps, Trees: Critical Responses to Franco Moretti,* edited by Jonathan Goodwin and John Holbo, 49-55. Anderson: Parlor Press, 2011.

Sculley, D., and Bradley M. Pasanek. 2008. "Meaning and Mining: The Impact of Implicit Assumptions in Data-Mining for the Humanities." *Literary and Linguistic Computing* 23, no 4: 409-424.

Sears, John. 2011. *Stephen King's Gothic*. Chicago: University of Chicago Press.

Stableford, Brian. 2004. "The Emergence of Science Fiction" and "Science Fiction Between the Wars." In *Anatomy of Wonder,* edited by Neil Barron, 5th edition, 3-44. Westport: Libraries Unlimited.

Suvin, Darko. 1972. "On the Poetics of the Science Fiction Genre." *College English* 34, no. 3: 372-82.

The First Hundred Years of Detective Fiction, 1841-1941. 1973. Lilly Library, Bloomington, IN. http://www.indiana.edu/~liblilly/etexts/detective/

Underwood, Ted. 2014. "Understanding Genre in a Collection of a Million Volumes." Figshare. https://dx.doi.org/10.6084/m9.figshare.1281251

Underwood, Ted. 2016. Code and data to support "The Life Cycles of Genres." Zenodo DOI to be arranged. For right now, in a repo here: https://github.com/tedunderwood/fiction

Wickham, Hadley. *ggplot2: elegant graphics for data analysis* New York: Springer.

Wolfe, Gary K. 2011. *Evaporating Genres: Essays on Fantastic Literature.* Middletown: Wesleyan University Press.

Appendix A: Metadata.

The metadata used in modeling is available in the github repo as finalmeta.csv. Each work can bear an unlimited number of "genre tags" characterizing different groups it is associated with. This table explains the meaning of the tags; there are a total of 962 texts. Most of the texts are volume-sized, but a few are short stories.

tag	#texts	dates	description or source
det100	89	1829 - 1941	*The First Hundred Years of Detective Fiction, 1841-1941.* 1973. Lilly Library, Bloomington, IN. http://www.indiana.edu/~liblilly/etexts/detective/
chimyst	146	1923-1989	Works categorized by librarians as "detective" or "mystery fiction," collected at the Chicago Text Lab.
locdetmyst	45	1832-1922	Works categorized by librarians as "detective and mystery fiction," collected in HathiTrust.
locdetective	16	1865-1912	Works categorized by librarians with the subject heading "Detectives." Often casebook fiction.
crime	2	1972-1974	Works categorized by librarians as "crime fiction" but not "detective fiction."
cozy	10	1920-1952	Works by authors mentioned as writing country-house mysteries in *The Mystery Readers' Advisory: The Librarian's Clues to Murder and Mayhem*, by John Charles, Joanna Morrison, and Candace Clark (Chicago: ALA, 2002).
hardboiled	10	1929-1970	Appendix to Geoffrey O'Brien, *Hardboiled America* (New York: Van Nostrand Reinhold, 1981).
newgate	7	1828-1840	Keith Hollingsworth, *The Newgate Novel,* Detroit 1963.
sensation	14	1860-1880	"The Sensation Novel," Winifred Hughes, in *A Companion to the Victorian Novel,* ed. Patrick Brantlinger and William B. Thesing (Blackwell, 2002).
lockandkey	10	1800-1903	Works anthologized in *The Lock and Key Library: Classic Mystery and Detective Stories,* edited by Julian Hawthorne

			(New York, 1909). Includes writers like Dostoevsky who probably were not seen as writers of mysteries even in 1909; not included in the article's model of "detective fiction."
pbgothic	96	1764-1988	"Chronology" in David Punter and Glennis Byron, *The Gothic* (Malden: Blackwell, 2004).
stangothic	21	1791-1834	A small subset of works tagged as "Gothic" in Stanford Literary Lab metadata.
lochorror	4	1818	Works tagged as "horror" by librarians, collected in HathiTrust.
chihorror	23	1933-1989	Works tagged as "horror" by librarians, collected in the Chicago Text Lab.
locghost	28	1826-1922	Works tagged as "ghost stories" by librarians.
locscifi	21	1836-1909	Works tagged as "science fiction" by librarians and collected in HathiTrust.
chiscifi	144	1901-1989	Works tagged as "science fiction" by librarians and collected at the Chicago Text Lab.
femscifi	9	1818-1922	Ockerbloom, Mary Mark. 2015. "Pre-1950 Utopias and Science Fiction by Women."
anatscifi	36	1771-1922	Stableford, Brian. 2004. "The Emergence of Science Fiction" and "Science Fiction Between the Wars." In *Anatomy of Wonder,* edited by Neil Barron, 5th edition, 3-44.
chiutopia	13	1920-1976	Works tagged as "Utopias" by librarians, collected at the Chicago Text Lab, not folded into "science fiction" in this article.
chifantasy	53	1901-1989	Works tagged as "fantastic" or "fantasy fiction" by librarians, not folded into "science fiction" for the purposes of this article.
---	---	---	---
juvenile	23	1904-	Works for a juvenile audience; collected but not used

		1922	in this article.
drop	33	1838-1922	Works that I decided not to use, left in the metadata for transparency. The most common reason is that they are juvenile.
random	169	1769-1922	Works randomly selected from HathiTrust Digital Library, using fiction metadata developed in the NEH-funded project "Understanding Genre in a Collectiono f a Million Volumes." "Random selection" here means that the volumes were selected randomly but then approved or rejected by the author, to avoid stray volumes of nonfiction, classical poetry, juvenile works, etc.
chirandom	202	1920-1989	Works randomly selected from the Chicago Text Lab. Selection here was more genuinely random. Note that both "random" tags can coexist with other genre tags. A randomly-selected volume could also be "chimyst," for instance; in that case it will be excluded from the negative (contrast) set only if "chimyst" is in the positive set.
teamred	484	1760-1989	Randomly selected authors for a sanity check.
teamblack	500	1764-1989	Randomly selected authors for a sanity check.
stew	224	1764-1989	A random selection of volumes balanced between Gothic, science fiction, and crime/detective traditions, in order to create a ghastly genre stew.

Appendix B.

Is the github repo containing code, data, and metadata for the project, located here:

https://github.com/tedunderwood/fiction