Abstract
How do we compare eliterature forms? What does it mean for a work to be implemented as hypertext, interactive fiction, or chatbot? "Benchmark fiction" is a methodology for creating 'benchmarks' - sets of adaptations of the “same” eliterature content across different media for the purpose of comparative study. While total equivalence between the resulting 'benchfic' is impossible, praxis remains important: by creating 'equivalent' media and then critiquing them, we revealing our own definitions of media through process. Work on the first story to be benchmarked, “The Lady or the Tiger” (1882) by Frank R. Stockton, inspired a framework for displaying sources through interchangeable display modules. The project is considered in terms of historical precedents (Lorem Ipsum, Hello World, Cloak of Darkness, Gabriella Infinita), contemporary theories (adaptation, remediation, media-specific analysis, transmedial and cross-media storytelling), and current experiments (chatbots, wikis, search art, cellular automata), with some discussion of design and pedagogy.

1. INTRODUCTION
A benchmark fiction, or ‘benchfic,’ is an elit adaptation for the purposes of comparing media. The term ‘benchmark’ here is playfully repurposed from the fields of computer science and strategic management in order to emphasize the focus on utility and standards. While ‘benchmark’ originated as a surveying term for a point of reference, in contemporary computer science, ‘benchmarking’ has come to mean the execution of a software test in order to ascertain the relative performance of underlying hardware.

As creative theoretical practitioners, we approach the benchmark test with painters’ smocks instead of lab coats. Nonetheless, the comparison to computer science benchmarking is serious. The production and comparison of benchfic is analogous to this process of testing hardware via software. With benchfic, the soft ‘content’ of a story might be ‘run’ on the hard ‘form’ of two different systems of elit implementation in order to examine differences in those two specific forms. Thus, Frank R. Stockton’s short story “The Lady, or the Tiger” might be adapted as a Storyspace hypertext fiction as well as an Inform interactive fiction. Alternately, the content of a work of elit (e.g. “Afternoon: a story’ by Michael Joyce”) might be separated from that form (Storyspace) and then re-implemented on some other form (HTML).

The underlying tenet of producing benchfic – the separation of ‘form’ from ‘content’ – is highly problematic. It is difficult or impossible to pinpoint where form ends and content begins in a given work, particularly artistic work designed to be experienced as a unified whole. Yet, with benchfic as with other processes of adaptation, translation and remediation, the problem of determining which elements to hold constant and which to vary is in large part the value of the undertaking.

In the process of producing benchfic, one’s concept of ‘form’ is formalized, as one’s vision of the content takes new shape. These very formalizations may break under the weight of the creative experiments, testing their limitations, another goal of the Project. There is no one correct approach. Rather, the Benchmark Fiction Project proposes the ongoing aggregation of a multitude of parallel elit adaptations; each with their own claim to what in translation studies is termed ‘equivalence.’ Rather than creating simple assumptions about how elit forms operate, benchmark fiction creates an opportunity to critically examine the assumptions and arguments we already make.

This investigation is both process and product oriented. First, through the process of attempting adaptation into target forms, we hope to identify and share different writing practices that emerge while trying to render similar effects in different new media. Second, we hope to generate example products which allow critical communities more opportunities to make apples-to-apples comparisons of new media fiction implementations – similar matter can be pulled up in two different applications and read closely side by side.

Ultimately, these experiments may prove more interpretive than empirical, in that we will not constrain ourselves to attempting pedantic remediations of the text into each form, but rather adapt the text in order to accentuate the strengths and limitations of each form. The results will reflect particular readings of the source text as well as interpretations of the nature of the forms themselves.

2. BENCHFIC IN THEORY
2.1 Adaptation
Adaptation studies are concerned with relations between texts, in particular, what passes from one text to another. As Kamilla Elliott so aptly observes, adaptation is a “heresy” for it “suggests that form is separable from content” [6]. In an attempt to call for a “renewed scrutiny” of the study of the relations between literature and film adaptations, Elliott provides a concept explication of adaptation studies. She outlines six approaches to understanding form and content within a book and film adaptation paradigm:
These are Psychic; Ventriloquist; Genetic; De(re)composing; Incarnational and Trumping.

The Psychic Concept of adaptation has two levels: the idea that an adaptation must preserve the “spirit of the text” (which doesn’t mean a high degree of sameness, but a text that persists certain perceived authorial intentions) and the idea that the spirit of a text passes from the text to adapter, to the adaptation and then to the reader or viewer. Unlike the Psychic Concept, a Ventriloquist approach sees “what passes from novel to film in adaptation [as] a dead corpse rather than a living spirit”. The film breathes life into the “dead” novel. The Genetic view sees the transitory substance as being an “underlying “deep” narrative structure akin to genetic structure”. A story, for instance, is the same, but with a different plot. A De(re)composing approach values a de- and re-composition of the text: elements in the film for example “serve to fulfill the disappointed hopes and desires of its characters”. An Incarnation view of adaptation sees the film as a necessary materialization that is prefigured, indeed demanded, by the novel. And finally, the Trumping relation between a novel and film is that of competition. In this approach, novel and film are compared according to which “represents better”. Within the conceptual paradigm of such concerns, the Benchmark Fiction Project intends to explore elit-adaptations for the sake of “renewed scrutiny” of this important area.

2.2 Transmedial Narrative and Game

For a given text there are both media-specific and transmedial qualities, and benchmarking seeks to explore their difference. In this regard, our research hypothesis is akin to narratologist David Herman’s:

Although narratives in different media exploit a common stock of narrative design principles, they exploit them in different, media-specific ways, or, rather, in a certain range of ways determined by the properties of each medium [12].

However, we extend this hypothesis to include ludic design concerns such as system, players, artificial conflict, rules, quantifiable outcome. These “key elements” were put forward by Katie Salen and Eric Zimmerman as follows:

A game is a system, players interact with the system, a game is an instance of conflict, the conflict in games is artificial, rules limit player behavior and define the game, and every game has a quantifiable outcome or goal. [26]

The aims, therefore, are to investigate the properties of the medium that “determine” the narrative and ludic features, and the nature of the effect. Some elit types, for instance, as considered more literary than ludic, and so the adaptations will be on two levels: at the level of genre and the level of mode of interaction.

2.3 Cross-media Storytelling

Beyond observing the hierarchical relations between texts is another approach: that of considering the sum of all the texts as a work. This approach is within the domain of “cross-media storytelling”, which has many antecedents: Kristeva’s [16] and Genette’s [10] “intertextuality”, Bakhtin’s “dialogism” and “heteroglossia” [2], Deleuze and Guattari’s “assemblage” [4], Foucault’s notion of a “work” [9] and Richard Wagner’s ’gesamtkunstwerk’ or total work of art [28], but was popularized by media theorist Henry Jenkins. In 2001, Jenkins observed the occurrence of increasing media channels and the introduction of stories that are delivered over multiple channels [13]. Through an analysis of “The Matrix” franchise, Jenkins posited the notion of “transmedia storytelling”:

In the ideal form of transmedia storytelling, each medium does what it does best—so that a story might be introduced in a film, expanded through television, novels, and comics, and its world might be explored and experienced through game play. Each franchise entry needs to be self-contained enough to enable autonomous consumption. That is, you don’t need to have seen the film to enjoy the game and vice-versa.

While talk of “franchise” and “autonomous consumption” reflects the capitalistic model of mass entertainment, Neo’s ability to take shape in the outside world and in the Matrix world speak directly to our notion of translation. More recently, Jill Walker has explored the notion of “distributed narrative”, where a narrative is distributed over time, space and producers [29]; while Lizbeth Klastrup and Susana Tosca have investigated the qualities of a work that can persist over multiple channels and time through the notion of “transmedial worlds” [17]. Semiotician Jay Lempke has also recognized the critical approaches needed to capture the peculiar instances of multi-text environments:

Not only do we not have adequate models of semiotic effects and inter-discursivity for each of these media individually, but many of the discursive and ideological effects of interest in inter-media franchises depend on inter-relations among presentations in coordinated, multiple semiotic media. [18]

All of these approaches require the ‘text’ to be viewed in a multi-text, and multi-channel environment, what Christy Dena conceives as “polymorphic works: narrative in many forms” [5]. Polymorphic works are an extension of literary theorist Itamar Even-Zohar’s “polysystem studies”, a theory addressing the complex socio-semiotic phenomena surrounding translations [7]. They place works within a continuum of multi-text relations ranging from repurposing to “transfiction” (story fragments distributed over different texts). Beyond, yet inclusive of, adaptation, a multi-text approach to storytelling extends the “de(re)composing” approach observed by Elliott.

The pervading principle of polymorphic works is that stories are systems rather than texts available in a single location, single-point-in-time by a singular author. They are experienced within a narrative universe in which the work is the sum of a variety of texts with varying relations. Each text plays a vital role in the work, creating the “work” rather than supplementing or subverting. Within such a framing any group of texts can be a “work”. In an article on “transfictionality” Marie-Laure Ryan addresses this issue by introducing conditions under which transfictionality can be identified: texts must be distinct and must be considered fictional, worlds must be distinct, and there must be an assumption of familiarity and an intentional preservation of transfictionality. These investigations illustrate audience and academic perceptions of entertainment as existing within a multi-text paradigm, situating Benchfic practice as a contemporary and highly applicable pedagogical tool.

Benchmarking enters this discourse as a creative and critical approach that recognizes the phenomenon of cross-media
storytelling and assists in the development of new media by providing a concentrated collection of texts with various inter-text relations.

2.4 Media Specific Analysis

As adaptations, a benchfic must vary while retaining some equivalence with the source text. A set of benchfic is adapted in parallel from a single source, and thus, by extension, must both vary from and have some equivalence with each other. What is the relationship of these specific differences to this general commonality?

In “Writing Machines,” N. Katherine Hayles describes Media Specific Analysis (MSA) as a mode of critical interrogation alert to the ways in which the medium constructs the work and the work constructs the medium... MSA attends both to the specificity of form...and to citations and imitations of one medium in another. MSA moves from the language of text to a more precise vocabulary of screen and page, digital program and analogue interface, code and link, mutable image and durable mark, computer and book. [11]

Characteristically of Hayles’ work, the emphasis here is on the specifics of materiality, rather than transcendent terms such as work or text.

Our analysis works towards this more precise vocabulary, even while it treats a work transcendentally. We propose that not only does the physical medium define the final product, but that the specific software has an influence over defining the parameters of all works that are produced using it. A simple example might be the use of Storyspace, the hypertext authoring and publishing system developed by Jay Bolter, John Smith, and Michael Joyce, then published by Eastgate Systems. Joyce recounts wanting to develop a system that would allow him to tell a story in a way that the codex book could not (personal interview). Later hypertexts, such as Shelley Jackson’s “Patchwork Girl” (1995) explore the limits of that system, but remain expressions of the capacities and constraints present in the software package. Macromedia Flash presents another example. Flash movies vary widely in their characteristics, but they all exist as expressions of what Flash permits.

This view of elit authoring as existing within a certain scope of available formal experimentation recalls Lev Manovich’s claim that “the greatest avant-garde film is software such as Final Cut Pro or After Effects” [20]. Such a piece of software, in Manovich’s terms, “contains the possibilities to combining together thousands of separate tracks into a single movie, as well as setting various relationships between all these different tracks – and thus it develops the avant-garde idea of film as an abstract visual score to its logical end.” We take this notion further, to explore the ways in which pieces of authoring software come to shape their fields, or perhaps, to use Deleuze and Guattari’s model, plateaus of art objects.

At some point it becomes difficult to separate the software from the material systems that run them, and we do not want to gloss over the differences in user experiences based on the machines they use to view these systems. Media Specific Analysis warns against believing in an easy equivalence that floats above material difference. For our purposes, where we differ is in our greater emphasis on the idea of representation and structure in digital text (e.g. the HTML H1 tag) as opposed to the screen instance (e.g. 24pt Arial Bold). This balances the specificity of presentation with the specificity of the code.

2.5 Remediation

In “Remediation,” Jay David Bolter and Richard Grusin discuss how media oscillate between transparent immediacy and opaque hypermediacy. While the process of mediation is ever-present, immediacy is a rhetoric that naturalizes the media frame and effaces it from view, while hypermediacy emphasizes the technical qualities of the medium and holds them forward for greater attention. The stances of natural being vs. skillful overcoming become an agonistic game, which is in particular played out as the New Media child tries to kill the Old Media father, and become him. In particular, their definition of a medium as “that which remediates” [3] and remediation as “representation of one medium in another” help us to see how that individual media forms are understood through difference from one another as they enter the mediasphere - in particular, digital media, for which remediation is a “defining characteristic”. In the case of elit media forms, wikifiction is for example that which is not hypertext fiction in specific ways - that these differences exist is certain, but what they are remains for us to elucidate.

Bolter and Grusin outline four degrees of remediation: transparency, translucency, hypermediacy, and absorption. In the first degree, new media becomes a container (“transparent”) to provide access to the older medium, a representation that is presented “without apparent irony or critique”. In the second, the old medium is represented faithfully, but the “electronic version is offered as an improvement”. No longer transparent, the new medium is “translucent”. Moving from such unobtrusive intentions, the third attempts to “refashion the older medium or media entirely, while still marking the presence of the older media”. The two become a media collage of sorts in which “hypermediacy” is the dominant aesthetic. Similarly, the final degree of remediation involves the new medium trying to “absorb the older medium entirely”, indeed, trying to replace it. However, as opposed to invention, the dependency on the older medium locks the old and new mediums together in remediation.

In these terms, the act of benchmarking is a move towards a more opaque hypermediacy. Benchfic are not creations that are to be accepted on their own terms, with attention only to whatever natural inner logic that might entail. Rather, benchmarking is a comparative process, and when an act of adaptation evokes what in translation studies would be termed a 'target' text in terms of some 'source,' the rhetoric of derivation and construction de-naturalizes the target in the eyes of the viewer. The resulting benchfic becomes opaque and significant in formal feature to the viewer, rather than a transparent conduit of contents.

Benchmarking is in part a practice of construction adaptations, and in part a critical methodology of attending to them for the purposes of media comparison. Interestingly, benchfic criticism also disrupts not only rhetorics of transparent immediacy, but also the progression towards absorption. Instead, it presents a kind of arrested hypermediacy. This is because benchfic are deeply and fundamentally comparative, yet in a way without priority. As only one potential 'target' among many, any benchmarked media is always contemplated in the context of peer media, interrupting remediation's vertical rhetorics of lineage of homage with a more horizontal perspective.
3. BENCHFIC IN HISTORY
Has their been anything like benchfic before? As the emphasis of remediation on modes of reception reminds us, any set of works including an original and several adaptations might be benchfic when considered together for the purpose of comparing media – regardless of how that set was created. Yet identifying previous creation or use communities is difficult. Here are a few possible examples.

While benchfic may be a mode of understanding, benchmarking can be described as an act and a creative process. What other practices or traditions have used creative variation for the purposes of evaluation? Two initial examples of such praxis arise from ancestors to eliterature – specifically, the fields of graphic design and software engineering.

3.1 Lorem Ipsum
One longstanding tradition in publishing and design is the use of invariant placeholder content. This passage is poured by graphic designers into sample layouts in order to give them the appearance of being filled without providing actual content to distract from the ‘form’ of the design (layout, typography, etc.).

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

In use since the 1500s, the lorem ipsum passage has probably appeared in a greater variety of presentations and formats than any passage in the history of written language. However, rather than interacting with those forms by constraining or being constrained by them, the purpose of the passage was to fill without filling, and to appear readable without being read. Illiteracy in Latin only aids in the desired effect, conveying an impression of legibility can be evaluated without the eyes being inadvertently drawn from the surface of the page into the meaning of the text. Interestingly, this led to a situation in which print culture became the steward of a text that was commonly believed to be a nonsensical mish-mash of pseudo-Latin. In fact, this common knowledge was incorrect. Richard McClinton of Hampden-Sydney College identified the lorem ipsum passage as an excerpt from Cicero’s 45 BC treaty on ethics “de Finibus Bonorum et Malorum.”

Latin was scrambled, it became as incomprehensible as Greek; the phrase ‘it’s Greek to me’ and ‘greeking’ have common semantic roots! [21]

Throughout its miraculous voyage down through the ages, the lorem ipsum passage has highlighted the media specific qualities of a given design through a particular kind of invisibility or transparency. It achieved this through a combination of naturalized convention and customary illegibility. Lorem Ipsum is not structured content (with an outline, footnotes, etc.), which interacts with the constraints of a particular design, but rather unstructured content, which fills it. Rather than being read in difference contexts, it is that which is not read at all. As such it represents one limit case on what benchfic can be – for it is reading that is being benchmarked, and without reading, the formula of common content in disparate forms can have no meaning.

3.2 Hello, World!
In computer programming there exists a tradition of writing a minimal program, which simply outputs text to some display. By convention, this text is always “Hello, world!,” and thus a vast catalog of Hello World programs (as they are called) have made their way into textbooks, manuals, and documentation for every imaginable computer language.

Each example is in fact two texts. The first, the output or result, is held invariant (“Hello, world!”), in order to highlight the various syntactical means whereby the various programming languages achieve the end of similar behavior using the second, the source code or cause. At present, numerous catalogs collect these program source codes side-by-side for mutual consideration—the Wikipedia article “Hello World Program” currently lists 153 and counting:

While small test programs existed since the development of programmable computers, the tradition of using the phrase “Hello world!” as the test message was influenced by an example program in the book The C Programming Language, by Brian Kernighan and Dennis Ritchie, published in 1978. The example program from that book prints “hello, world” (i.e., no capital letters, no exclamation sign; those entered the tradition later). The book had inherited the program from a 1974 Bell Laboratories internal memorandum by Kernighan —Programming in C: A Tutorial— which shows the first known version of the program.

Although the 30-year voyage of Hello World is less impressive than Lorem Ipsum’s 500 (or 2050) year journey, both were circulated by agents largely ignorant of their origins. Both examples are not merely text, but the combination of a conventional text with a cultural practice within a creative community. Both practices emphasize the explication of design syntax – in the case of Hello World, this design syntax being the programming language itself.

The difference between evaluating the layout and the code however is the difference between the generally ‘framing’ or ‘surface’ status of the layout and the generally ‘underlying’ or ‘prior’ status of the code. Whereas in graphic design evaluation takes place by considering the holistic page (without being distracted by the meaning of words, which exist within that page),
in programming evaluation proceeds more along the lines of Norbert Weiner’s information theory, by testing the channel of message passing (which must emerge out of or because of the compiled code). Thus, while the text “Hello, world!” is as much pseudo-content as “lorem ipsum,” the practice of Hello World brackets the question of reading differently. Rather than suspending reading, it essentializes reading by presenting a minimal-length message which is equivalent to a binary return – a “yes” or “1” that responds to the unspoken question “did the program work, or didn’t it?”

3.3 Adventure and Cloak of Darkness

A practice of benchfic in elit, however, would not just be an examination of outputs or ends, but an examination of models of interaction. If there is a precedent to benchfic in software, it probably begins the ur-text of interactive fiction (IF), “Adventure,” a text whose history is described by IF scholar Nick Montfort as intimately tied to the rise of the internet, computer culture, and the game industry [14]. Co-authored and widely adapted / reimplemented from its inception, “Adventure” later entered a new stage of ubiquity when it was repackaged as the commercial Zork series by Infocom and became the landmark text of the emerging commercial computer game industry.

The repackaging of “Adventure” arose out of a culture of reimplementation that was already widespread, however the interest here is in Infocom’s strategy of distributing their interactive fictions via a virtual machine, abstracted from that first text and subsequently implemented on countless hardware platforms.

Like later virtual machines (such as Java), a z-code interpreter could execute instructions on any platform that met the specification. Whereas the specification Hello World is almost the simplest imaginable, the z-code interpreter was quite complex, and specifying it represented to some extent a formalization of the possibilities of the genre of IF as imagined by Infocom.

Today the distribution of IF virtual machines is among other things a cultural tradition in programming and specifically within the community of UNIX/POSIX distributions, where its inclusion in an operating system is seen as a point of historical pride.

With IF interpreters however, much like the significantly simpler “Hello World,” the goal was not to examine variations in behavior, but rather portability, achieving nearly identical behavior by means of widely varying code. Today, even though simple z-code interpreters are available as everything from command-line applications to Firefox browser plug-ins, their behavior remains largely the same.

However it is perhaps not surprising that IF’s culture of portability gave rise to another project with the specific purpose of providing a kind of benchmark or comparative metric. Roger Firth’s “Cloak of Darkness” website implements the same short scenario about a mysterious message through a variety of interactive fiction languages, with some discussion of the varying capabilities and limits of each language. [8]

This site tries to help in your evaluation, by presenting the same (very small) game using a range of authoring systems. The implementations have been made reasonably consistent, so as to facilitate comparison. As well as the game source... we sometimes provide information on how it was compiled, present a transcript showing it being run, and try to mention some real games that you might also like to try.

As with Hello World the emphasis in “Cloak of Darkness” is on achieving similar behavior – yet unlike it the purpose here is also to reveal dissimilarities of effect (reader experience) rather than simply cause (author implemented code). In this Cloak seems like a synthesis of these two practices, and a more sophisticated model of evaluating variation in interactive media. However, substantial differences from the concept of benchfic remain. The target audience for Cloak is primarily authors looking to evaluate potential languages for use in future projects, rather than critical readers considering the reading / playing experience in varying contexts. In addition, Cloak conceives of itself as operating within a genre (IF) rather than across genres (IF, hypertext fiction, chatbots, etc.). To that end a ‘specification’ – a descriptive list of component features to be included in the scenario – rather than a source, something that must perhaps be conceived of more flexibly in order to operate across disparate genres. A retelling of the same Cloak scenario in widely disparate forms would expand the comparative focus from IF’s varieties to its limits, highlighting what is characteristic of IF and what lies beyond.

3.4 Gabriella Infinita

Whereas “Cloak of Darkness” is our most sophisticated example of a collection of adaptations proceeding from a specification towards a concept, other sets of adaptations have been the work of a single author, proceeding from concept into specificity.

One useful example is Jaime Alejandro Rodríguez Ruiz’s “Gabriella Infinita” [24] a set of texts that are available as ebook (PDF), hypertext, and hypermedia. For Rodríguez Ruiz, a professor of electronic literature at the Pontifical Universidad Javeriana in Bogotá, the various intermedia iterations of the tale follow his own education in hypertext and hypermedia. He refers to the entire production as an “obra metamorifico” or “metamorphic work,” and indeed the story has had various formal incarnations. He began with a novel, or as he puts it, “It began as a bud, greedy about its body, realizing its fragility and contingency and in the end realizing that it was destined for volatility.” (Of course, this translation of Rodríguez Ruiz’s words also possesses such volatility.) Indeed “Gabriella Infinita” seems to yearn for an infinite, boundless state. The novel begins with Gabriella arriving at the scene of her now missing lover, Frederico, a tormented genius author. On that note of the author gone missing, the text uses multiple points of view, metafictional twists, and all the self-reflexivity and narrative complexity of the works of postmodernism. It is a book very much at odds with or in contest with its form. In 1997, Rodriguez Ruiz encountered hypertext as a literary form at a conference on the novel and postmodernity. As his book seemed to be yearning for this new form, he adapted the work. Following the current of technology, he adapted the hypertext to hypermedia with the collaboration of voice actors, visual artist Clara Inés Silva, and programmer Carlos Roberto Torres.

Although Rodriguez Ruiz presents his works as moving “towards” hypermedia, by posting all of the forms with commentary, his work begs for comparative analysis. In many ways, the fixed sequence text form of the story presents more versatility than the later forms. Its pages can be accessed in any order, its voices and images supplied by the user. Also, as a PDF, the full text is searchable. On the other hand, the addition of film and voice clips
adds layers of information and intonation that cannot be found in the print-only text. In any case, this work in its current online form stands as the archeological strata of cities built upon each other, retaining basic infrastructures, while adding new technology. Though presented in a narrative of progress, these proto-benchfics offer valuable lessons in creative adaptations and revisions in different electronic forms.

4. BENCHFIC IN PEDAGOGY

4.1 Hypertextual Writing

Stuart Moulthrop created another early predecessor to benchfic, when he adapted Borges's "Garden of Forking Paths" for the Internet. Then, in a first year college writing course at Carnegie Mellon University, he and Nancy asked students to Moulthrop's hypertext. Moulthrop and Kaplan argue that "with hypertext, the range of options broadens, allowing narratives that at least approximate Yu's vision of infinite pathways" [23]. One student appended additional nodes to Moulthrop's piece. In turn, they interpreted the student's creative action, arguing, "In the space of hypertextual writing, anything that arises will be merged, gathered into the network of polyvalent discourses." On the one hand, here is an example of students learning to express interpretations through creative electronic adaptations. Adapting a print text into hypertext performs an interpretation of the text through the medium of hypertext as a critical lens and electronic form. On the other hand, their problems with copyright lead us to choose a work that was out of copyright, which also fits our open source aesthetic.

4.2 Hypertextual Reading

Empirical reader response investigations into how the reading process changes with the hypertextual text are highly relevant to the proposed application of the Benchfic Project. Studies have been undertaken to compare print and hypertextual reading processes, but also compare different hypertextual designs. Imaginative hyperlinking may for instance, be at odds with conservative reading practices. In a case study conducted by David S. Miiall and Theresa Dobson, a text, Elizabeth Bowen's "The Demon Lover", was delivered in a "structurally linear format" (next buttons only) to one group and to another in a "simulated hypertextual form" (hyperlinked words) [19]. The story showed that the readers of greater hypertextual form took longer, felt confused, focused on the mechanics of the reading process, and difficulty following the narrative but enjoyed "control" over the plot and increased suspense. Such results help hypertext creators to understand at what point their creativity inhibits effective communication and the experience, assists in understanding the rhetoric of hypertextual constructions and how reading changes in different navigation strategies. Indeed, continuing on with the empirical reader response tradition, the Benchfic provides students with a collection of texts with different hypertextual strategies (as employed with different elit types) to analyze according to rhetorical and reading concerns.

5. BENCHFIC IN PRACTICE

5.1 Our First Source: The Lady, or the Tiger?

When selecting a first text for the Benchmark Fiction Project to experiment with, we had a number of criteria. To enrich the introspective process of creation, we hoped for a story whose themes resonated with new media or elitertext. To make the process of producing benchfics accessible, we wanted a piece that was short, yet to keep them rich we wanted one that was thematically substantial. A range of characters, times, locations, thoughts, and physical actions were desirable, but again that range needed to be limited. Finally, we wanted one that was free and clear of all the sundry complications of international copyright law - and, if possible, available in some widely disseminated electronic text form such as via the Guttenberg project. In practice, this meant either published under a recent and explicit copyleft license (such as Creative Commons) or else something so old as to have lapsed into the public domain. Previous elit adaptations, such as Stuart Moulthrop’s “Forking Paths: An Interaction after Jorge Luis Borges,” had to be disassembled after just such complications.

In the history of the short story, “The Lady, or the Tiger?” (LOTT) by Frank Richard Stockton stands as a kind of anomaly [27]. Written in 1882, it is the tale of a “semi-barbaric” kingdom with a uniquely quantum system of justice. Any alleged criminal is brought into the arena and presented with two doors. Behind one waits a beautiful woman whom he will marry (as all criminals are men, presumably). Behind the other paces a hungry and ferocious tiger. On a bed or on a plate, justice is served.

Unfortunately, for the princess, who is also “semi-barbaric,” her lover, having been caught in a crime, is the next to stand trial. She learns what is behind each door, and when her accused lover looks up to her before choosing, she gestures to the right. But we do not know that she has gestured towards the door with the woman, because the Princess knows the lady behind the door and suspects that she is no lady. To be sure, the Princess suspects that that woman has had designs on her lover. But how could she feed her lover to a tiger? The story ends with the author declining to resolve the question, and asking the reader to choose the outcome. We are left in the world of the indeterminate.

Instructors in writing workshops attempt to stave off a tide of similar open-ended stories by telling their students “there can only be one ‘The Lady, or the Tiger.’” Nonetheless, Stockton’s indeterminacy prefigures the rise of a post-modernism aesthetic that eschews resolution, now a hallmark of contemporary fiction.

LOTT is also unusual in its final request that the reader intervene as writer – a second person mode that is the hallmark of both the multilinear “Choose Your Own Adventure” gamebooks and IF. As an early exemplar of the indeterminate story, LOTT is often seen as a precursor to the multiple paths realized in later multilinear storytelling. Indeed, hypertext author Michael Joyce wrote LOTT into hypertext history by referencing it in his seminal “afternoon: a story.” [15] His lexia labeled “The Lady or the Tiger” teases:

It comes down to that, doesn’t it? Despite what we think of our techno-philosophical advancement? Love or death. Risk with two faces. Go on, press the button, treat it all as if it were real. The lady? or the tiger?

Choosing the hyperlink of “the lady” however sends the reader to a lexia labeled, “you have no choice.” For Joyce, invocation of LOTT at the inception of a new genre (literary hypertext) echoes the novel rhetoric of the original, while, simultaneously questioning any illusions we have of “techno-philosophical advancement’" – just as the king believes himself to be enacting true justice through a spectacle of terrible symmetry, our belief in the progress of storytelling via interactive electronic forms may a romantic (or barbaric) fantasy.
5.2 Benchmarking as Database

Our source selected, the Benchmark Fiction Project begin experimenting with various elit forms for rendering the text. Though relatively new to media history, electronic literature has produced a large number of genres and subgenres, with even more media-specific channels for producing those genres. To offer a brief, but by no means all-inclusive list, we might consider “The Lady, or the Tiger” via: Interactive Fiction (Inform, TADS), Hypertext (Storyspace, HTML), Hypermedia (Flash, PowerPoint), Chatbots (AIML, Yapanda, Personality Forge)… the list becomes even longer the more specificity we include in an implementation. What version number of the Inform language? AIML implemented via JavaScript over the web, or rather over an instant messaging network?

Our decision was to subsume this general discussion into a lightweight framework for contributions—a database cataloging system that could serve as a prototype and visual metaphor for thinking through our critical practice. Our move to database modeling follows Lev Manovich’s thesis in “The Language of New Media” that the database is the essential conceptual form of new media art, and performs benchfic as a kind of meta-art—a new media creation in its own right. Designing the catalog led to developing the following project vocabulary:

**Source:** A source is simple digital text with minimal markup—a file with a simple notational structure. A given unit might be a literal raw text, or be made up of subordinated units such as pages, paragraphs, chunks, lexias, or various adaptations such as a dialog, a lexicon, etc. A given unit might be identical or almost identical across various sources, and we can compare differences. We currently have sources of “The Lady, or the Tiger?” which are broken up by paragraph, annotated with wikicode, and parsed as word-pairs, among others.

**Interface:** An interface is a system for rendering content of a particular type in a particular form. It may be a print text, wiki, CYOA, chatbot etc. In conception, this might be compared to other functional form / content systems such as CSS / HTML. In the practice, the interface behaves as a function that takes a source and returns an edition, and as such includes not just display but also some kind of navigation or interaction model that might head into the AJAX end of what is possible via HTML, or beyond. We already list several PHP displays, a CGI script, and a few remote web APIs. Interfaces are in principle reusable—by dividing the source from the interface, the benchmarker makes a particular assertion about where the essence of the text ends and the essence of the form begins. A raw text module or a wiki module might take the same source collection of paragraphs and render them differently. Not only can an interface render many sources, but one source might be rendered through several different interfaces. We currently have two such alternate wiki display engines, which draw on the catalog database, as well as interfaces for browsable paged text, printable versions, a Pandorabot, and some whimsical external examples that make use of programs such as GoogleFight and Life vs. Life.

**Edition:** An edition is a specific pairing of a benchmark source with an interface—any interface can display many sources, and a source may be displayable by many interfaces. While most editions are dynamically generated, editions may also be external, or even static. The possibility of static editions in the Benchfic Catalog is particularly important, as it means that bringing an adaptation into consideration next to other benchfic does not require following this methodology of sources and interfaces. Allowance for listing external sources also means that strong centralization of the catalog is not required. Contributions that follow such a methodology but cannot be subsumed under our current database schema can likewise list their results in the catalog. Furthermore, the current technical specification of Linux-Apache-MySQL-PHP (LAMP) is a free, widely supported standard for open source and academic projects, which should ease dissemination of Benchmark Fiction software when it reaches a stage suitable for distributed educational use.

5.3 Wikifiction: TheLady or TheTiger?

Given our various research interests, the most interesting interface-editions to our project members are larger adaptations for hypertext fiction, chatbot fiction, and interactive fiction. These are large undertakings, however, and at this early stage we are beginning with smaller undertaking as proof-of-concept.

Perhaps the most complete benchfic to date is the wikifiction interface used to render “TheLady or TheTiger?” Wikifiction is a branching text much like hypertext fiction, yet taking advantage of the peculiarity of wiki syntax that the bracketed word like [this] or “CamelCase” word LikeThis is automatically rendered into a link. Such a link leads to a new chunk or lexia whose title is the same as the link text.

While in modern wikis such as Mediawiki (which runs Wikipedia) this constraint can be removed through the use of extended syntax, the principle that “links lead to named objects” is a foundational one for the wiki, and creates an interesting aesthetic for rendering forking or garden-path fiction through a wiki interface.

For this wikifiction interface, we chose TiddlyWiki, a DHTML wiki display script which represents the wiki entries as floating boxes which can be reshuffled by opening or closing. The result is a reading effect not unlike what Ted Nelson described as “stretchtext.” The source preserves a default order for reading the text by “naming” each paragraph after a wikified word or phrase in the previous paragraph, thus rendering the text as a chain or as an accordion.

For the LOTT benchfic done via this interface, one of the facts this highlighted was the extent to which each paragraph does not logically queue the next one, creating a need for mid-paragraph
The introduction to the conversation reads:

Behind the door lies a lady, pacing the floor in anticipation or a tiger, cleaning its teeth for its close up. The king has offered you the chance to pass notes scribbled on scraps back and forth between you and one of the doors. The king’s mage has cast a spell allowing the tiger, for the moment, the powers of speech. From your interactions and questions, you must decide who is behind this door. The Lady or the Tiger...

http://www.pandorabots.com/pandora/talk?botid=f92e9ba343e5ea30

The user can then attempt to discern which one is behind the door. Asking, “What is your gender,” yields the reply, “My gender is Tigress.” Asking, “What color is your hair,” elicits, “My hair is striped.” The playful ambiguity of the answers returns to the indeterminacy of the story and its infamous ending. The adaptation uses the situation of the questions to challenge the interactor’s position. Asking, “Are you a lady,” elicits “You wouldn’t ask a lady a question like that.” The piece also refers to its own quantum state. Asking about the King elicits, “The King, who made this system of divine justice, making the captive choose between Schrödinger and the Cat, taught his daughter every thing she knows.” With the answers, the chatbot becomes more of a complement to and a commentary on the tale than a strict re-presentation.

This chatbot is an AIML Alicebot, using a customized version of Richard Wallace’s 2002 A.L.I.C.E. code and hosted on the Pandorabots site. Much of A.L.I.C.E.’s responses have been left, and, more importantly, A.L.I.C.E.’s response algorithm has not been changed except in the routine customization of responses. This proves to be one chatbot edition of LOTT and other chatbots, whether from scratch or using another authorware system, such as UltraHal, would have to be benchmarked separately.

5.4 Chatbots: Turing or the Tiger

The choice of how to adapt the chatbot to the tale is almost as perilous as the choice within the LOTT itself. Should the chatbot tell the story when questioned? Should the chatbot be one of the characters (e.g. the Princess or the King)? This case underscores a fundamental question in benchmarking fiction: How much license should the author take when adapting the story to new media forms that are structurally different (not based on linear print text) from the original, particularly from the standpoint of reader/user/viewer interaction?

To date we have dealt with this question by allowing ourselves the maximum freedom within each form, realizing that merely using one medium like another does not do the testing justice. Just as early filmmakers learned that merely positioning a camera before a dramatic scene was not the best use of the camera (a movie is in fact not a play seen from the best seat in the house), new media artists explore the unique properties of the medium rather than simply remediating. For example, a chatbot could tell the entire story in one response to any given input, but that would not be using the chatbot’s functionality. Nonetheless, we realize that this freedom may contaminate our benchmarking process on some literal level that is frankly outside of our aesthetic interests in new media. In effect, we have added many more doors in the process of adapting LOTT.

In one adaptation of a chatbot to LOTT, the text on the HTML page presents the user with a variation on the prisoner’s dilemma.

The user can then attempt to discern which one is behind the door. Asking, “What is your gender,” yields the reply, “My gender is Tigress.” Asking, “What color is your hair,” elicits, “My hair is striped.” The playful ambiguity of the answers returns to the indeterminacy of the story and its infamous ending. The adaptation uses the situation of the questions to challenge the interactor’s position. Asking, “Are you a lady,” elicits “You wouldn’t ask a lady a question like that.” The piece also refers to its own quantum state. Asking about the King elicits, “The King, who made this system of divine justice, making the captive choose between Schrödinger and the Cat, taught his daughter every thing she knows.” With the answers, the chatbot becomes more of a complement to and a commentary on the tale than a strict re-presentation.

This chatbot is an AIML Alicebot, using a customized version of Richard Wallace’s 2002 A.L.I.C.E. code and hosted on the Pandorabots site. Much of A.L.I.C.E.’s responses have been left, and, more importantly, A.L.I.C.E.’s response algorithm has not been changed except in the routine customization of responses. This proves to be one chatbot edition of LOTT and other chatbots, whether from scratch or using another authorware system, such as UltraHal, would have to be benchmarked separately.

5.5 Google Fight: Lady vs. Tiger

Our interpretations of the “source” of LOTT are more playful in the Google Fight and Life vs. Life interfaces. These experiments push at the limits of what can be considered a legitimate adaptation, or benchfic.

Google Fight is a website that uses the Google search engine to pit two search terms against each other as if in a physical battle. After the terms have been entered, a brief flash animation plays of stick figures battling. Google is searched for each term, and the resulting page counts are represented as bar graphs, with the more commonly matched search term crowned as winner.

As a first attempt at feeding LOTT through an external interface, we took significant word pairs from the story and fed them into the Google search engine to see the result. When confronted with the choice between “The Lady” or “The Tiger,” what would the PageRanked internet find collectively more interesting?

Our “Lady Vs. the Tiger” edition of LOTT is that particular fight. More precisely, the edition is a link leading to a live (re)enactment. Each user witnesses a result that may be similar or different, as the search results will fluctuate throughout the life of the site. As of 6/23/05 at 12:24 pm, PST, the Lady was ahead by a slight margin (34,200,000) to the Tiger’s (33,500,000). After watching these battles, we might question whether LOTT was really about a choice between the Lady and the Tiger. What if it was between Trust and Suspicion? As of this writing, Trust wins by a landslide. Left Door vs. Right Door? King vs. Princess?

Google Fight allows for variability in each instantiation and keeps the story in flux. It also complicates the idea of the binary choice by resolving it through an enormous quantitative calculation, itself the result of a qualitative algorithm (PageRank). Of course,
GoogleFight might work better integrated into a more contextual adaptation of the tale, as the two search terms offer no back-story, exposition, or narrative outside of combat. Nonetheless, the central indeterminacy of the story as well as an accentuation of chance have been brought out in this whimsical adaptation.

5.6 Life vs. Life: The Lady regiT ehT
If GoogleFight resolves the final chapter of LOTT through leveraging the aggregated digital text of the web, Life vs. Life is an algorithm for resolving conflict between individual pixels.

Life vs. Life is a website which provides a head-to-head rendition of the Life cellular automata algorithm. In it, a red and a blue pattern compete, and, like in the game Othello, the color with the most dots wins when no more moves can be made. The system is set up with a pattern authoring environment and a competition rankings board wherein any two patterns may compete. With a fixed grid and a limit of 60 units per pattern, a few simple aesthetics of warrior-patterns have emerged over time to dominate the rankings.

What if we treated the pixel-fonts of “The Lady, or the Tiger?” as material, and tried to resolve the conflict operationally? Rather than design killer geometric forms, the interface is here repurposed to run bitmapped text fonts, which allows the rules of Life to determine which of two textual propositions overcomes the other – which text propagates to success, and which text fails and is consumed. Life vs. Life is drive by the metaphor of bacteria, which are not only reproducing and competing (as in the original Life) but also consuming one another. “The Lady” may defeat “The Tiger” in these examples, but she does so by becoming that which consumes. The situation here mimics that of the princess in the original story – a woman about to watch her lover consumed by one of two possible fates, either of which will separate them forever.

There are many other interesting qualities of the medium (mirroring, determinism, the performance of the battle over time), however many of the hard decisions in creating a benchfic relate to pragmatic decisions over choosing font point sizes, deciding on visual arrangement such as spacing and alignment, dividing up the text (what to do with the “or”?!) and so on. The life algorithm is extremely sensitivity to initial conditions, making the composition process an odd echo of the precarious nature of the lover’s choice, whose fate balances on a knife’s edge.

Our Life Vs. Life adaptation plays with the contemporary new media interrogations of “the computational universe,” an epistemology based on Ed Fredkin’s infinitesimal calculations of cellular automata produce the world. Through this adaptation, we see LOTT enacted on the level of subatomic particles, or a universe determined by countless subatomic doors.

6. COMPARISONS
Our early results prove some interesting lessons about LOTT. One outcome has been the evolution of our understanding of the story. As with any form of adaptation, the original must be understood in terms of what we determine to be essential elements. Through both literalist transcriptions and fanciful elaborations on those elements, we gain a greater sense of the stakes of the text. Yet we also come to appreciate what kinds of stashes are inherent in the forms of the benchfic. The energy of the insight lies in synergy.

Indeterminacy plays out differently in the different forms. In wikifiction, the topic headings highlight the changing stake of the conversation. In GoogleFight, the seemingly random evolution of the networked pages of the internet come to substitute for the operations of chance or Fate. In Life vs. Life, our intervention in initial conditions propels us into the butterfly-effect unfolding a determinate universe, while in chatbots, that indeterminacy is the question of the unknown entity behind the interface and the play of the guessing game of Turing’s “Imitation Game.” While this last example does not directly match the situation of LOTT, it does put the user in the same position of the interrogator in a situation where (in Turing’s proposal) choosing the wrong door may diminish his/her own life. The choice of the Tiger may very well match mistaking the man for a woman, or a computer for a human. It is a problem of ontology that by merely being posed can be fatal to a world view.

Already this reading begins to move away from the analysis of form and back towards the analysis of the content of the story. This reciprocal relationship will be fruitful. Further, the story has helped point out some of the essential characteristics of the form, such as the unknowability of the entire internet, or the experience of risk and chance when Googling, or the computational universe that underlies Life Vs. Life. That this commentary seems to be a kind of meta-analysis, involving layering on significance to the forms, also distinguishes this process from benchmarking in the computer science sense. Benchmarking Fiction is a creative and interpretive activity that will yield different results to different heuristics, even when examining the same edition.

7. FUTURE DIRECTIONS
Even the concept of Benchmark Fiction can be adapted, and our open source cataloging system encourages as much. The works will be accessible through a website, along with essays on their creation. The goal is enabling others to upload their own creative and theoretical contributions. Hopefully, Benchmark Fiction will serve as an initial methodology towards systematic evaluation and productive exploration of the electronic forms available to authors of electronic literature – although others would be welcome. No doubt, choosing other texts, other forms, or even other implementations of those forms would yield valuable results. Further, these results can serve as lessons to theorists and practitioners of new media who wish to understand the limitations and strengths of the various forms. In our example, benchmarking is applied creatively, as each adaptation reinterpreted this story. As an example of and call for a kind of creative practice, benchmarking and benchfic may also prove to be a source for re-examination of the story and new media itself, much in the way
adaptation and revision have operated in oral, print, and other artistic milieu for centuries.

8. REFERENCES