Research

Peer Review:

Copyright:

Open Access:

Digital Preservation:

RESEARCH

Theme Enrichment Analysis: A Statistical Test for Identifying Significantly Enriched Themes in a List of Stories with an Application to the Star Trek Television Franchise

Mikael Onsjö[1] and Paul Sheridan[2]

[1] Independent Researcher, GB

[2] Hirosaki University, JP

Corresponding author: Paul Sheridan (paul.sheridan.stats@gmail.com)

In this paper, we describe how the hypergeometric test can be used to determine whether a given theme of interest occurs in a storyset at a frequency more than would be expected by chance. By a storyset we mean simply a list of stories defined according to a common attribute (e.g., author, movement, period). The test works roughly as follows: Given a background storyset and a sub-storyset of interest, the test determines whether a given theme is over-represented in the sub-storyset, based on comparing the proportions of stories in the sub-storyset and background storyset featuring the theme. A storyset is said to be "enriched" for a theme with respect to a particular background storyset, when the theme is identified as being significantly over-represented by the test. Furthermore, we introduce here a toy dataset consisting of 280 manually themed Star Trek television franchise episodes. As a proof of concept, we use the hypergeometric test to analyze the Star Trek stories for enriched themes. The hypergeometric testing approach to theme enrichment analysis is implemented for the Star Trek thematic dataset in the R package stoRy. A related R Shiny web application can be found at https://github.com/theme-ontology/shiny-apps.

Keywords: enrichment analysis; hypergeometric test; over-representation analysis; Star Trek; theme ontology

Dans cet article, nous décrivons la manière dont le test hypergéométrique peut être employé pour déterminer si un thème d'intérêt donné se produit dans une gamme d'histoires à une fréquence qui est plus élevée que ce à quoi on s'attendrait avec le hasard. Par gamme d'histoires, nous entendons simplement une liste d'histoires regroupées selon un attribut commun (par exemple : auteur, mouvement, période). Le test fonctionne à peu près

comme ceci : Après avoir reçu une gamme d'histoires en contexte et une sous-gamme d'histoires d'intérêt, le test détermine si le thème donné est surreprésenté dans la sous-gamme d'histoires, en comparant le taux d'histoires dans la sous-gamme d'histoires et dans la gamme d'histoires en contexte qui présentent le thème en question. Une gamme d'histoires est considérée comme « enrichie » pour un thème par rapport à une gamme d'histoires en contexte particulière lorsque le test identifie le thème comme notablement surreprésenté. Par ailleurs, nous présentons ici une base de données de jouets qui se compose de 280 épisodes de la franchise Star Trek qui ont été manuellement organisés par thème. Comme preuve de concept, nous nous servons du test hypergéométrique pour analyser les histoires de Star Trek avec des thèmes enrichis. La méthode de test hypergéométrique pour l'analyse de thèmes enrichis est réalisée pour la base de données thématique de Star Trek dans le logiciel R stoRy. Une application Web similaire de R Shiny se trouve sur le site-Web suivant : https://github.com/theme-ontology/shiny-apps.

Introduction

A *literary theme*, or *theme* for short, is loosely defined as "An idea that recurs in or pervades a work of art or literature" (Lexico 2019). Themes are often expressible in a single word or short phrase, as is illustrated by such garden-variety themes as "love", "loyalty" and "the lust for gold". These examples happen to be value-neutral abstractions. But themes may just as well take the form of morally charged messages, such as "be wary of strangers" and "do not judge prematurely". The consummate story-maker usually takes pains to imply a theme indirectly, rather than state it explicitly. Sometimes the story-maker is even unconscious of important themes found in their stories. A typical story will feature multiple themes. In the present work, we distinguish between *central* themes (i.e., themes found to recur throughout a major part of a story or otherwise important to its conclusion) and *peripheral* themes (i.e., briefly featured themes that are not part of the main story narrative). To sum up: the themes of a story furnish the partaker thereof with an executive summary of what the story is truly about.

Enrichment (or over-representation) analysis, which amounts to a class of statistical tests, is well-suited for determining whether a given theme of interest is featured more than would be expected by chance in a subset of stories drawn from a larger set of stories. The hypergeometric test uses the hypergeometric distribution to calculate a *p*-value associated with having drawn at least k successes in n draws, without replacement, from a population of size N that contains exactly K successes. We define a *storyset* to be a set of stories defined by some common attribute (e.g., author, movement, period). In the present context, it is appropriate to think of the population as a *background storyset* (i.e., a set of N stories), and the sub-population as a *test storyset* (i.e., a subset of n ≤ N background stories). A test storyset is said to be *enriched* (or *over-represented*) for a given theme when the associated hypergeometric test *p*-value is lower than a predetermined significance level.

The hypergeometric test is routinely employed in bioinformatical analyses to identify over-represented biological terms in lists of genes (Boyle et al. 2004; Zheng and Wang 2008; Huang et al. 2009). This differs markedly from the corresponding state of affairs in literary studies. To our knowledge, the hypergeometric test is yet to be implemented in any of the textual analysis tools in common usage among humanities scholars, including the Stanford Topic Modeling Toolbox (Ramage et al. 2009), TAPoR (TAPoR 3 2019), TOME (Klein et al. 2015), Word Seer (Muralidharan et al. 2013) and Voyant Tools (Sinclair et al. 2019). The same holds true of computer-assisted qualitative data analysis software that is sometimes used for document analysis in the social sciences (What Is Qualitative Research? 2019). ATLAS.ti (Muhr 2004), NVivo (QSR International Pty Ltd. 2015) and MAXQDA (VERBI Software 1989-2019) are three such programs that allow users to manage, analyze, and visualize data related to text, audio, and video documents. However, none of these programs implement the said test at the time of writing.

We champion the hypergeometric test as a statistically principled approach to theme enrichment analysis. To this end, we introduce a toy dataset consisting of 280 thematically annotated Star Trek television series episodes. We manually annotated

each of the 80 episodes of *Star Trek: The Original Series* (*TOS*), 22 episodes of *Star Trek: The Animated Series* (*TAS*), and 178 episodes of *Star Trek: The Next Generation* (*TNG*) with themes drawn from the Literary Theme Ontology (LTO) (Sheridan et al. 2019). The LTO is to our knowledge the only non-corpus specific ontology of literary themes. The thematically annotated Star Trek episode dataset paves the way for a demonstration of how the hypergeometric testing approach to enriched theme identification is helpful for summing up what makes a storyset unique and for generating speculative hypotheses. We present the results of two case studies as a proof of concept. The first examines how the portrayal of the Klingons changed from that of a tyrannical, expansionist empire in *TOS/TAS* to that of an inward-looking warrior culture in *TNG*. The second explores the topmost significantly enriched themes for each of *TOS*, *TAS* and *TNG*. In short, we find that *TOS* stands out for its focus on the important social issues of the day, *TAS* for novel sci-fi and fantasy concepts especially suited to an animated series and *TNG* for its comparatively refined treatment of the human condition. In addition, we show that our results compare favorably with those obtained using the standard term frequency–inverse document frequency (TF-IDF) approach to enrichment analysis (Salton et al. 1975), which is commonly implemented in textual analysis software packages. More specifically, we find that the hypergeometric testing and TF-IDF approaches to enrichment analysis, while agreeing in broad outlines, are liable to exhibit conspicuous differences in terms of the topmost enriched themes they identify.

In light of these findings, we recommend the hypergeometric test (which yields standard probabilities) over the commonly used TF-IDF heuristic (which yields non-standard numeric scores) for thematic enrichment analyses. Moreover, interpretation is facilitated by virtue of the test operating within a conventional statistical testing framework. For instance, a hypergeometric test p-value of 0.05 for a theme X means there is a 1 in 20 chance of theme X being featured in at least k out of n stories selected at random and without replacement from among N background stories. TF-IDF numerical scores are not so easily amenable to interpretation. However, we stress that the TF-IDF heuristic is perfectly appropriate for common information

retrieval tasks, such as the ranking of document relevance in response to a user query, where it is used as a ranking function.

The rest of the paper is organized as follows: In Section 2 we provide an overview of Star Trek film and television series franchise. In Section 3 we introduce the LTO version 0.1.1. It is a hierarchically organized controlled vocabulary of themes, partitioned into the following four domains: the human condition, society, the pursuit of knowledge and alternate reality. There are 1535 unique themes in total. Criteria and guidelines motivating our hierarchical arrangement are discussed. In Section 4 we explain the hypergeometric testing approach to theme enrichment analysis in full technical detail. In Section 5 we use the hypergeometric test to identify enriched themes in the two Star Trek case studies outlined above. These results we compare with those obtained using the standard TF-IDF approach to enrichment analysis. We conclude the paper in Section 6 with a summary of our main contributions, a discussion of some limitations of our methodology and go on to describe a handful of possible future directions. Most notably, in terms of limitations, we emphasize that we manually annotated Star Trek episodes with themes, and as a consequence the findings we report inevitably reflect our point of view, and are not fully replicable. The theme enrichment analysis procedure based on the hypergeometric test is implemented in the R package stoRy (version 0.1.1) (Sheridan and Onsjö 2017), released through CRAN (The Comprehensive R Archive Network 2019). The thematically annotated Star Trek episode dataset is included in the package. A related R Shiny web application is available for download at the Theme Ontology GitHub repository (Theme Ontology Project GitHub Repository 2019).

The Star Trek television series and film franchise

Star Trek has long been an important part of American popular culture (Haynes, 2016), and remains popular among sci-fi enthusiasts the world over (STARFLEET 2019). Those acquainted with the Star Trek science fiction media franchise will know that in the main, it comprises seven television series and thirteen feature films (Star Trek 2019). **Table 1** shows an overview. The first series, which is known as Star Trek: The Original

Table 1: Star Trek television series overview.

TV Series Title	Short Name	Original Release	No. of Seasons	No. of Episodes
Star Trek: The Original Series	*TOS*	1966–1969	3	79
Star Trek: The Animated Series	*TAS*	1973–1974	2	22
Star Trek: The Next Generation	*TNG*	1987–1994	7	178
Star Trek: Deep Space Nine	*DS9*	1993–1999	7	177
Star Trek: Voyager	*Voyager*	1995–2001	7	172
Star Trek: Enterprise	*Enterprise*	2001–2005	4	99
Star Trek: Discovery	*Discovery*	2017–present	2	27

Series (or simply *TOS*), began airing in 1966 and ran for 80 mostly delightful episodes. It depicts the adventures of Captain James T. Kirk and his crew aboard the starship Enterprise on a five-year mission to explore the galaxy. The mission was temporarily suspended three years in when the show was cancelled in 1969 on account of poor ratings. But the Enterprise crew saw their mission through to completion in Star Trek: The Animated Series (*TAS*), which ran from 1973 to 1974 in two seasons consisting of 22 somewhat less delightful episodes. Six feature films following the *TOS/TAS* cast on subsequent adventures were released in the years from 1979 to 1991. From *TOS* was spawned the spin-off television series Star Trek: The Next Generation (*TNG*) which ran from 1987 to 1994 in seven seasons consisting of 178 episodes. It is set a generation or so after Captain Kirk's five-year mission. In the series, a fresh cast of characters is led by Captain Jean-Luc Picard on a similar mission of galactic exploration aboard a newfangled starship Enterprise. There are four associated feature films. Four subsequent television series have been produced: Star Trek: Deep Space Nine (1993 – 99), Star Trek: Voyager (1995 – 2001) and Star Trek: Enterprise (2001 – 05). Three reboot films based on *TOS* have also been released to date and a seventh television series, Star Trek: Discovery (2017 – present), is presently being aired.

The literary theme ontology

A theme ontology is a controlled vocabulary of defined terms representing literary themes in fiction. In this section, we describe the LTO version 0.1.1. It is draft ontology

of literary themes covering a broad range of recurring themes in fiction (Sheridan et al. 2019). An LTO literary theme is taken to be either a topic that is featured in a work of fiction or an opinion that is conveyed about a topic featured in a work of fiction (Sheridan et al. 2019). In the jargon of literary theory, this means that LTO encompasses both literary subjects and literary themes (Griffith 2010). LTO upper-level organization is inspired by the traditional theme classification system of literary critic William Henry Hudson (Hudson 1913). The version of LTO described here consists of 1535 unique themes arranged into the following four upper-level domains:

The Human Condition [●]: Themes pertaining to "characteristics, key events and situations which compose the essentials of human existence, such as birth, growth, emotionality, aspiration, conflict and mortality" (Wikipedia 2019).

Society [●]: Themes pertaining to a "community of people living in a particular country or region and having shared customs, laws and organizations" (Lexico 2019).

The Pursuit of Knowledge [●]: Themes pertaining to "facts, information and skills acquired through experience or education; the theoretical or practical understanding of a subject" (Lexico 2019).

Alternate Reality [●]: Themes related to subject matter falling outside of reality as it is presently understood. These are classical science fiction and fantasy themes (The Science Fiction Encyclopedia 2019).

The LTO follows Basic Formal Ontology (Smith and Bittner 2005) best practices (Arp et al. 2015), and is encoded using Web Ontology Language (OWL2) (Hitzler et al. 2012). LTO version 0.1.1 is available for download in OWL format at the Theme Ontology GitHub repository under a Creative Commons Attribution 4.0 International license (CC BY 4.0) (Sheridan and Onsjö 2019A). We encourage non-technical users to explore a more recent developmental version of the LTO at the Theme Ontology website (Theme Ontology 2019).

Figure 1 shows a bird's eye view of the ontology. The abstract theme "literary thematic entity" is taken as root theme. Each domain is structured as a tree descended
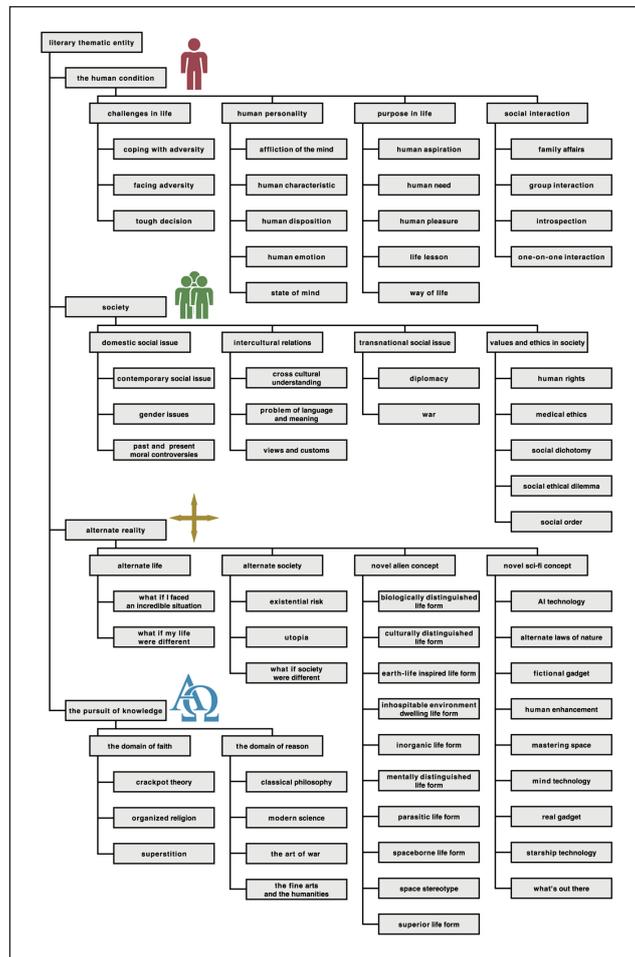
Figure 1: Literary Theme Ontology version 0.1.1 domain hierarchies shown to three levels of depth.

from the root with "the human condition", "society", "the pursuit of knowledge" and "alternate reality", serving as the top themes of their respective domains. Each child theme is made to bear a subtype relationship with its parent. In the figure, the ontology tree structure is depicted to a height of three levels, although in reality, it branches out a number of levels further still, as summarized along with other information in **Table 2**.

The scope of LTO constitutes any and all precisely definable literary subjects and literary themes that can be expected to recur in multiple works of fiction. In designing

Table 2: Literary Theme Ontology version 0.1.1 domain hierarchy summary statistics.

Domain Root Theme	Domain Color-code	Theme Count	Leaf Theme Count	Tree Height
the human condition	🔴	631	578	6
society	🟢	279	256	4
the pursuit of knowledge	🔵	235	217	4
alternate reality	🟡🟡	390	357	4

the ontology, we strive to define sibling themes so as to be mutually exclusive, but not necessarily jointly exhaustive. All themes are accompanied by short definitions in an effort to make their range of applicability plain. We appeal to the principle of refutability as an anodyne to vagueness in definition writing. In other words, a well-defined theme will be such that there is the possibility of appealing to the definition to show that the associated theme is not featured in a story. Take "the desire for vengeance" as an example, which is defined as "A character seeks retribution over a perceived injury or wrong.". We contend that definition writing of this sort helps to bring the conversation of whether a theme is featured in a given story into the realm of rational argumentation.

We populated the LTO with terms by watching all *TOS*, *TAS* and *TNG* Star Trek television series episodes, of which there are 280, and documenting the themes as we saw them. We deemed Star Trek to be useful for this purpose on account that it is a culturally significant franchise which treats a wide range of human nature, society and speculative fiction themes. We assigned central and peripheral themes for each episode. Note that centrality and peripherality are to be understood not as properties of a theme, but rather as relations between a theme and a story. Thus, it is perfectly legitimate for a given theme to be central to one story and peripheral to another. A theme is said to be *observed* in a story when it is featured in such a way that none of its descendent themes are also featured. An ancestor of an observed theme in a story is said to be a *latent* theme. **Table 3** shows a basic statistical summary of the thematic data by series. Summary statistics for observed themes are shown in the upper part of the table, and observed and latent themes in the lower part. While the

Table 3: Star Trek television franchise series episode theme summary statistics.

Series ID	Episode Count	Mean number of central themes per episode ± s.d.	Mean number of peripheral themes per episode ± s.d.
Observed Themes			
TOS	80	12.54 ± 4.40	20.04 ± 6.28
TAS	22	6.68 ± 2.55	3.50 ± 2.39
TNG	178	11.72 ± 4.42	14.93 ± 5.66
Observed and Latent Themes			
TOS	80	23.38 ± 5.61	29.38 ± 6.21
TAS	22	7.55 ± 2.06	5.23 ± 3.05
TNG	178	26.39 ± 7.99	30.72 ± 8.96

task of examining each of 280 Star Trek episodes for the presence of 1535 themes may strike the reader as a lifetime endeavor, the actual situation is not nearly so dire. In practice, once someone identifies an important topic in a story, they can quickly home in on those themes best representing the topic by navigating down the LTO hierarchy. For example, one may consider whether a story in question focuses an individual's experiences in life (i.e. "the human condition"). If so, then one may consider further whether the experience is about a single individual, two individuals interacting, or a small group of people interacting. Once the appropriate upper-level theme has been hit on, it is a small task to use the Theme Ontology website (Theme Ontology 2019) search function to peruse highly specialized themes to ascertain which best fits the situation at hand. The same sort of workflow applies to the other domains of "society", "the pursuit of knowledge" and "alternate reality".

Take the *TOS* episode *The Devil in the Dark* (1967) as an example. In the story, the starship USS Enterprise is dispatched to investigate rumors of a subterranean creature that is thought to be responsible for the destruction of equipment and the deaths of fifty men on the Janus VI mining colony. Captain James T. Kirk and First Officer Spock discover a hideous "silicon-based life form" inhabiting the surrounding bedrock. The mother Horta, as the creature comes to be known, is an "endangered species" – the

last of its kind. Kirk is faced with a "tough decision" as the creature seemingly blocks the miners path to wealth: either commit genocide or forego plundering the mother Horta's natural resources. But Spock fortuitously manages to achieve a "cross cultural understanding" with the creature by means of a Vulcan mind-meld. An unsettling compromise is reached when the mother Horta agrees to help the miners locate ore deposits in the rock in exchange for a cessation of hostilities. The attentive viewer will note, from the creature's perspective, an "attack from outer space by a powerful conquering alien race" (a.k.a. "alien invasion") with overtones of "the morality of colonization" by the episode's conclusion. This summarizes some of the more salient central story themes, as recorded by the authors. The proverb "beauty is in the eye of the beholder" is a noteworthy peripheral theme. Indeed, Spock learned, in the course of the mind-meld, that the human form is just as repellant to the mother Horta, as her appearance is to humans.

We annotated each of the 280 episodes of *TOS*, *TAS* and *TNG* with themes in a similar manner. The process according to which we assigned themes can be summed up as follows. We independently annotated episodes with themes and then compared notes with a view toward building a consensus set of themes for each episode. We aimed to abide in the principle of low-hanging fruit in the compilation of consensus themes. In the present context, this means we tried to ensure that at least the most salient topics featured in the episodes are covered by appropriate themes. Another principle guiding our thought process is the minimization of false positives (i.e., the tagging of episodes with themes that are not featured) at the expense of tolerating false negatives (i.e., neglecting to tag episodes with themes that they feature). This strategy amounts to erring on the side of caution. Particular theme usages are motivated with brief comments in an effort justify their applicability. In addition to providing a checkable written record of the episode annotations, the practice of writing justifications for themes helps to reduce the risk of annotating episodes with nonapplicable themes. We fully acknowledge that this process needs more safeguards against the annotating of stories with themes that are idiosyncratic and unique to our point of view. We will return to this subject in the discussion section.

Theme enrichment analysis

This section is devoted to an exposition on the hypergeometric testing approach to theme enrichment analysis. The test uses the *p*-value obtained from the hypergeometric cumulative distribution

$$P\left(k,n,K,N\right)=\sum\nolimits_{i=k}^{n}\frac{\binom{K}{i}\binom{N-K}{n-i}}{\binom{N}{n}}$$

to answer the question of whether a given theme occurs in a test storyset at a frequency significantly greater than would be expected by chance alone. In the equation, n is size of the test storyset, k is the number of stories in the test storyset featuring the theme, N is the size of the background storyset, and K is the number of stories in the background storyset featuring the theme. The size of a storyset is the number of stories it contains. The K background stories featuring the theme determine what we will call the *theme storyset*. **Figure 2** depicts the testing framework in Venn diagrammatic form. The value of P(k, n, K, N) is the probability of observing at least k stories featuring a given theme in a test storyset of size n that is composed

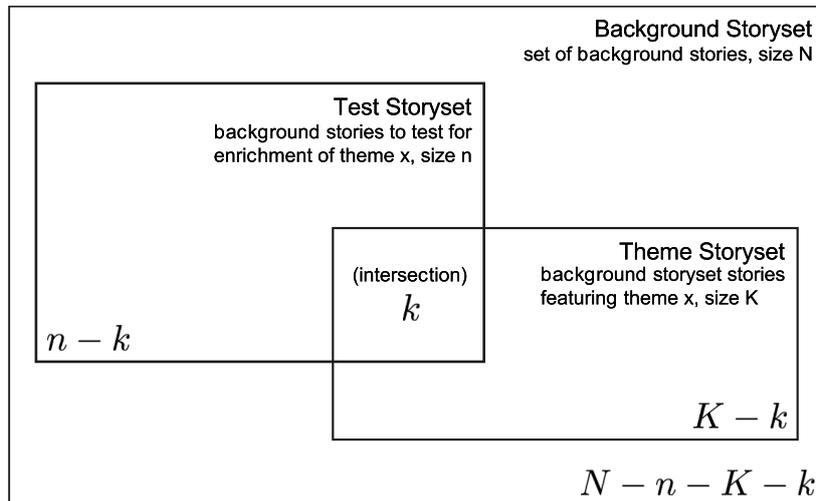
Figure 2: Hypergeometric testing framework overview for testing whether a given theme x is enriched in a test storyset relative to a background storyset. The various storysets and variables n, k, K, and N are defined in the main text.

of stories drawn at random, without replacement, from the background storyset. As we explained in the introduction, a theme is deemed to be enriched in a test storyset with respect to a background storyset if its *p*-value is less than a preselected significance level, alpha. Whether the test storyset is found to be enriched for a given theme or not will necessarily depend on the choice of background storyset.

The hypergeometric testing procedure is silent on the matter of theme, test storyset, and background storyset selection. It falls on the investigator to first pose an interesting question, and then choose the hypergeometric test input storysets accordingly. In the previous section, we proposed that the *TOS* episode *The Devil in the Dark* (1967) featured an "alien invasion". Indeed, the theme has long been a favorite among sci-fi writers. More generally, the sci-fi genre has proven to be particularly well-suited to the exploration of "existential risks" (i.e., manners by which civilization on a planetary scale and beyond could come into jeopardy), of which "alien invasion" is but one, albeit far-fetched, example. It is easy to imagine a fictitious investigator who wishes to assess whether "existential risk" (theme) is enriched in *TOS* (test storyset) with respect to all of *TOS*, *TAS* and *TNG* (background storyset). This casts the question "Does *TOS* stand out among Gene Roddenberry produced *Star Trek* television series for its featuring of existential risks?" in precise enough terms to be made amenable to analysis. As we will see in Section 4.2, the test results in a *p*-value of 0.0002, so that we may conclude the theme "existential risk" is enriched in *TOS* at significance level alpha equal to 0.05.

In practice, the curious investigator will often wish to check whether each theme in an ontology is significantly enriched in a given test storyset relative to some background storyset. This raises the specter of multiple comparisons (Noble 2009; Meijer and Goeman 2016; McDonald 2014). To illustrate the matter, recall that the *p*-value in a hypothesis test is the probability of getting a result at least as extreme as the observed one, assuming the null hypothesis is true. A *p*-value calculated to be less than the investigator's desired significance level, alpha, is interpreted as evidence in favor of the alternative hypothesis. The value of alpha sets an upper limit on the chance of making a false positive discovery that the investigator is willing to tolerate. Consider, for example, the classroom problem of testing whether a given coin is fair

or biased. The conventional null and alternative hypotheses are that the coin is fair and biased, respectively. Choosing an alpha of 0.05 translates into a 1 in 20 chance of concluding that the hypothetical coin in question is biased when it is actually fair (i.e., the null hypothesis is rejected when it is true). In the case when a large number of hypothesis tests are conducted, chance dictates that some definite proportion of the p-values obtained will be less than alpha, even when the truth of all the null hypotheses is assured. For example, suppose 1535 fair coins are individually tested in the above manner. This is as many coins as there are themes in our ontology. If all 1535 coins are tested at an alpha of 0.05, then the expected number of fair coins mistakenly identified as biased ones works out to be $0.05 \times 1535 \approx 77$. The aim of multiple comparison correction procedures is to limit the number of false positives. The Bonferroni correction and the false discovery rate (Benjamini and Hochberg 1995; Benjamini and Yekutieli 2001) are the two most common procedures. However, it is inappropriate to correct for multiple comparisons in the context of testing for multiple enriched themes. This is because a given theme is either over-represented in a test storyset relative to a background or it is not. If a theme is significantly enriched at a given significance level, then it is a true positive by definition. It would therefore be misguided to apply the sorts of standard correction procedures that are used in biology in the context of testing list of genes for enriched terms. For this reason, we refrain from correcting the p-values calculated in this paper for multiple comparisons.

A study of enriched themes in Star Trek

In this section, we present two cases studies of theme enrichment analysis applied to the Star Trek television franchise. Central and peripheral themes are included in the analyses. As we explained in the introduction, the first examines how the Klingons changed from a tyrannical and expansionist empire in *TOS* to an inward-looking warrior culture in *TNG*. The second, also described in the introduction, takes an in-depth look into significantly enriched themes by series. In short, we find that *TOS* stands out for its treatment of what constitutes a good society and how one is to lead a good life within it, *TAS* for novel alien and sci-fi concepts appropriate to

an animated series and *TNG* for its comparatively refined treatment of the human condition. We additionally present the reassuring results of some negative control experiments, and make the case that the hypergeometric test compares favorably to the TF-IDF approach to enrichment analysis. In particular, we find that while the overlap in enriched themes exhibited between the hypergeometric test and the TF-IDF heuristic results in mutually consistent interpretations, the two approaches nevertheless exhibit conspicuous differences. But the main message we wish to convey is that the hypergeometric testing approach to enrichment analysis is helpful when it comes to both the summarization what makes a storyset unique and the formulation of speculative hypotheses. We recognize the potential for confirmation bias to undermine the validity of our interpretations seeing as we both collected and analyzed the data. And we reiterate that our manual approach to thematic annotation is subjective to our point of view, and not fully replicable. However, while our experimental design leaves much room for improvement, we contend that it is not fatal to our goals of demonstrating the utility of the hypergeometric test for thematic enrichment analyses and evaluating it against TF-IDF. The enrichment analysis interpretations we propose are to be regarded as speculative hypothesis advanced in the way demonstrating proof of concept. Moreover, our favoring of the hypergeometric test over the TF-IDF heuristic for enrichment analyses rests on theoretical considerations that we use the data to exemplify. The thematic enrichment analyses are easily replicated using version 0.1.1 of the stoRy package.

A tale of two Klingons

The Klingons are an über-belligerent humanoid species in the Star Trek alien pantheon. In *TOS*, the Klingon Empire is made to pursue a harsh imperialist foreign policy, characterized by the unapologetic use of military force in the subjugation of their weaker neighbors. The Federation, by contrast, is portrayed as a group of confederated alien races, united under the common principles of human rights, equality, and interstellar cooperation. In the series, the Federation fights to check Klingon expansion in the galaxy. The conflict between the Federation and the Klingon Empire in *TOS* is commonly understood as an allegory for the Cold War

(Cantor 2000). According to this interpretation, the Federation represents the Western powers, and the Klingon Empire represents the Soviet Union; both as seen from a more or less contemporaneous American point of view. But by *TNG*, Klingon society had undergone a radical transformation. The Klingons went from being an avowed Federation enemy, intent on galactic domination, to being a loose Federation ally, preoccupied with internal struggle and the maintaining of antiquated warrior traditions (Knight and Smith 1998). The changing face Klingon society has been examined in much detail in books (Taylor 2002; Gonzalez 2015; Telotte 2008), academic papers (Knight and Smith 1998; Cantor 2000) and scattered online sources (BBC Online 2011; Roblin 2016; Bondurant 2017).

In this case study, we use the Klingons as a positive control to demonstrate that our approach to theme enrichment analysis is able to distinguish the imperialist Klingons of *TOS* from the warrior culture ones of *TNG*. To this end, we curated a storyset consisting of all 26 Klingon-centric *TOS*, *TAS*, and *TNG* episodes; see Supplementary Information File 1 for a complete episode listing (Sheridan and Onsjö 2019B). The criterion for episode inclusion is that the Klingons were deemed by the authors to have been featured throughout the episode in a way that is central to the story. We performed two theme enrichment analyses: 1) the test storyset of *TOS*/*TAS* Klingon-centric episodes against a background of all *TOS*/*TAS* episodes, 2) the test storyset of *TNG* Klingon-centric episodes against a background of all *TNG* episodes. The *TOS* and *TAS* episodes we pooled because *TAS* is conventionally considered to be a continuation of *TOS*. The *TOS*/*TAS* and *TNG* test storysets consist of n = 8 and n = 18 episodes, respectively. In each experiment, we calculated an enrichment score, that is, a hypergeometric test *p*-value, for each of the 1535 themes present in the ontology.

The test identified 27 and 22 enriched themes at significance level 0.05 in the *TOS*/*TAS* and *TNG* Klingon storysets, respectively. Tables containing the full results from these analyses are included in Supplementary Information File 2 (Sheridan and Onsjö 2019B). **Table 4** shows the top 20 most enriched themes for each analysis. We focus on interpreting the enriched themes shown in the table, but we note that the excluded ones are consistent with the interpretation we advance. With that

Table 4: Enriched themes in Klingon-centric episodes relative to *TOS/TAS* and *TNG* backgrounds, respectively.

Rank	Theme	Domain	k/K	P value	Comment
	Top 20 Enriched Themes in *TOS/TAS* Klingon Episodes				
1	über-belligerent alien	⬤	5/5	<0.0001	The Klingons are belligerence personified.
2	diplomatic negotiating	⬤	4/7	0.0005	Teetered on brink of war with the Federation.
3	culturally distinguished life form	⬤	6/20	0.0006	Parent theme of über-belligerent alien.
4	man vs. beast	⬤	3/5	0.0030	Beasts faced: mugato (×1), tribble (×2).
5	diplomacy	⬤	5/19	0.0053	Parent theme of diplomatic negotiating.
6	conflict over a shared resource	⬤	2/2	0.0054	Vied with the Federation over galactic resources.
7	atrocities of war	⬤	2/2	0.0054	Not above committing war crimes.
8	tribble	⬤	2/2	0.0054	Mortal enemy of the Klingons.
9	pacifism	⬤	3/7	0.0098	Pacifists made easy targets for conquest.
10	military tactics	⬤	3/7	0.0098	Masters of the art of war.
11	war	⬤	5/22	0.0110	Reveled in warfare and conquest.
12	transnational social issue	⬤	6/32	0.0111	Parent theme of diplomacy and war.

(Contd.)

Rank	Theme	Domain	k/K	P value	Comment
13	the art of war	🔵	5/23	0.0136	Parent theme of military tactics.
14	miscellaneous life form	🟡	3/8	0.0150	Parent theme of tribble.
15	imperialistic society	🟢	3/8	0.0150	Klingon Empire subjugated weaker neighbors.
16	conflict of moral codes	🟢	2/3	0.0157	Klingon imperialism vs. Federation benevolence.
17	cross cultural understanding	🟢	4/16	0.0198	Parent theme of conflict of moral codes.
18	humility	🔴	3/9	0.0217	Exhibited in the face of Klingon aggression.
19	patience	🔴	3/10	0.0297	Exhibited in the face of Klingon aggression.
20	temperance	🔴	2/4	0.0301	Contraposed with Klingon licentiousness.

Top 20 Enriched Themes in *TNG* Klingon Episodes

Rank	Theme	Domain	k/K	P value	Comment
1	über-belligerent alien	🟡	14/14	<0.0001	The Klingons are still belligerence personified.
2	honor	🔴	13/18	<0.0001	Central to the *TNG* Klingon way of life.
3	culturally distinguished life form	🟡	14/53	<0.0001	Parent theme of über-belligerent alien.
4	the need for cultural heritage	🔴	5/8	0.0003	Worf cherished his Klingon cultural heritage.
5	rage	🔴	7/23	0.0032	Flying into a violent rage is a key Klingon trait.

(Contd.)

Rank	Theme	Domain	k/K	P value	Comment
6	racism in society	🟢	3/5	0.0080	Worf despised the Romulan race.
7	facing wrongful accusations	🔴	3/5	0.0080	E.g. Worf falsely accused of treason.
8	belonging	🔴	7/27	0.0090	Worf was the lone Klingon in Starfleet.
9	religious fanaticism	🟢	2/2	0.0100	E.g. Klingon religious fanatics cloned their messiah.
10	war of succession	🟢	2/2	0.0100	Klingon civil war.
11	father and son	🔴	6/21	0.0101	E.g. Worf and his son Alexander.
12	guilt and evidence	🟢	4/11	0.0163	Applies by and large to non-Klingon characters.
13	the art of war	🔵	6/24	0.0207	Still masters of the art of war.
14	the lust for power	🔴	3/7	0.0244	Various Klingons lusted to rule Klingon Empire.
15	the desire for redemption	🔴	2/3	0.0280	Worf sought to redeem father's soiled reputation.
16	loyalty	🔴	7/33	0.0295	A ruthless yet loyal alien race.
17	agreeable characteristic	🔴	14/96	0.0305	Parent theme of honor.
18	brother and brother	🔴	3/8	0.0366	Worf and his brother Kurn.
19	cooperation	🔴	5/20	0.0370	Klingon Empire and Fed. sometimes joined forces.
20	surprise	🔴	6/27	0.0377	E.g. Worf was surprised to find he had a young son.

caveat aside, an inspection of **Table 4** shows that our positive control test results are interpretable in a manner that is in keeping with expectation. Consider first those themes from the society domain. *TOS/TAS* Klingon "imperialistic society" posed a serious military threat to the Federation. When an inevitable "transnational social issue" flared up between the Federation and Klingon Empire, such as a "conflict over a shared resource", the resolution usually came about by either "diplomatic negotiating" or outright "war". Although in rare instances, Klingons and Federation members came to a "cross cultural understanding" when united by a common enemy. But *TOS/TAS* Klingon society had at its heart a "conflict of moral codes" with Federation ideology that proved insuperable. Where *TOS/TAS* Klingon society is enterprising and enthusiastic in its convictions, *TNG* Klingon society is inward looking and gloomy. The society themes enriched in *TNG* Klingon episodes pertain to internal conflicts, as evidenced by the themes "racism in society", "religious fanaticism" and "war of succession". No longer is the Klingon Empire striving to impose Klingon values on the galaxy by means of military force, but rather it is focused on its own internal affairs. This brings us to the human condition. In *TOS/TAS*, the human condition domain themes are almost all virtues possessed by the aliens that the Klingons conquered (i.e., "pacifism", "humility", "patience" and "temperance"). On the other hand, the human condition domain themes enriched in *TNG* (i.e., "honor", "rage" and "loyalty") are all signature *TNG* Klingon characteristics. Notice, by contrast, that "honor" is nowhere to be found among the top 20 significantly enriched *TOS/TAS* Klingon themes. In fact, it occurs as the 101st ranked theme with a *p*-value of 0.219. What is more, in the one episode in which it was featured – *Friday's Child* (1967) – it pertained not to the Klingons, but to their enemy the Capellans. A number of the human condition themes pertain to Worf, who is the lone Klingon crew member aboard the Enterprise-D starship, trying to maintain his Klingon culture in a human world. The themes "the need for cultural heritage" and "belonging" are two noteworthy examples. Other human condition themes surround aspects of life in a cut-throat warrior culture, like "the lust for power" and "the desire for redemption". Finally, it is pleasant to note that Klingon "über-belligerence" and a passion for "the art of war" shine through in both cases.

A tale of three series

We used the test to identify enriched themes in *TOS* (120 themes), *TAS* (6 themes) and *TNG* (46 themes) at significance level 0.05. The background storyset in each case consists of the episodes from all the series combined. Here we report the outcomes of the analyses and show how they can aid in the generation of speculative hypotheses. In keeping with the Klingon case study from the previous subsection, we frame our hypotheses about *TOS* and *TNG* in terms of the top 20 most enriched themes for each respective series. This, we contend, is enough to convey the merits of our methodology without burdening the reader with protracted syntheses of long lists of enriched themes, however pleasant an exercise the formulation of such syntheses might be for the authors. On the other hand, we take some leeway and extend our analysis of *TAS* to the top 10 most enriched themes. But we note that our general conclusions would remain unchanged had we limited our interpretations to the 6 enriched themes at significance level 0.05. Tables containing the full results from these analyses are included in Supplementary Information File 2 (Sheridan and Onsjö 2019B). Star Trek enthusiasts will find few surprises in the kinds of themes that are shown to distinguish the respective series. To the layperson, however, the results of **Table 5** may be unexpected and serve as a useful point of departure for exploring the series. The stacked percentage bar plots of **Figure 3** show a broad pattern of human condition domain themes being enriched in *TNG*, alternate reality domain themes in *TAS* and society domain themes in *TOS* to some degree. The associated matrix scatterplot hints at some interesting enriched theme domain correlations between series. But let us proceed to inspect and compare more specific themes in order to gain a more nuanced understanding of the series.

TOS: The two society domain themes "female stereotype" and "gender issues" stand out as they relate to the role of women in 1960s society. The former is indicative of what Karen Blair has described as a tendency in *TOS* to portray females in such a manner as to "affirm traditional male fantasies in a most direct and unenlightened way" (Blair 1983). The latter, however, is in keeping with another line of scholarly thought that contends *TOS* made some positive contributions to the advancement

Table 5: Enriched themes for each Star Trek television series relative to a *TOS/TAS/ TNG* background.

Rank	Theme	Domain	k/K	P value	Comment
		Top 20 *TOS* Enriched Themes			
1	female stereotype	🟢	19/24	<0.0001	Reinforced some outmoded female stereotypes.
2	wrath	🔴	41/85	<0.0001	Violent outbursts abounded.
3	alternate society	🟡	45/100	<0.0001	Mind-openingly different societies explored.
4	facing a fight to the death	🔴	13/16	<0.0001	Captain Kirk no stranger to such altercations.
5	what if I had to fight to the death	🟡	13/16	<0.0001	Captain Kirk won ten and lost one to Spock.
6	gender issues	🟢	26/46	<0.0001	Challenged some outmoded sexist attitudes, too.
7	rage	🔴	27/50	<0.0001	Hurled bowls of Vulcan plomeek soup and so on.
8	real gadget	🟡	8/8	<0.0001	Flatscreen TV, lie detector, teleconferencing, etc.
9	unpleasant emotion	🔴	78/237	<0.0001	Parent theme of rage.

(Contd.)

Rank	Theme	Domain	k/K	P value	Comment
10	alternate life	🟡	39/87	<0.0001	Fantastical things befalling people explored.
11	way of life	🔴	24/45	0.0001	Ideas on how to lead a good life explored.
12	purpose in life	🔴	70/201	0.0002	Parent theme of way of life.
13	existential risk	🟡	37/83	0.0002	Various threats to human civilization explored.
14	imperialistic society	🟢	8/9	0.0002	Federation a bastion against imperialism.
15	what if I faced an incredible situation	🟡	30/64	0.0003	Examples of incredible situations faced are provided in the main text.
16	disagreeable characteristic	🔴	48/121	0.0003	E.g. complacency, deviousness, rudeness, etc.
17	personal ethical dilemma	🔴	45/112	0.0005	Characters faced with difficult moral choices.
18	tough decision	🔴	61/169	0.0005	Parent of personal ethical dilemma.

(Contd.)

Rank	Theme	Domain	k/K	P value	Comment
19	man-made existential risk	🟡	22/43	0.0006	E.g. the dangers of WMDs and societal laziness.
20	the need for a challenge in life	🔴	7/8	0.0008	Fundamental component of the *TOS* ethos.
Top 10 *TAS* Enriched Themes					
1	earth-life inspired life from	🟡	6/15	0.0004	Avians, felinoids, insectoids, slug-like aliens, etc.
2	life-support belt	🟡	3/3	0.0004	A belt-like device that functions as a spacesuit.
3	miscellaneous life form	🟡	5/13	0.0017	E.g. an alien composed of autonomous parts.
4	man vs. beast	🔴	3/8	0.0187	Beasts faced: ie-matya, rock beast, tribble.
5	Chariots of the Gods	🔵	2/4	0.0326	Aliens supplied the ancients with technology.
6	genetic engineering	🔵	2/4	0.0326	Captain Kirk genetically engineered to have gills.
7	what if my life were different	🟡	6/38	0.0615	E.g. what if I were a specimen animal in a zoo.

(Contd.)

Rank	Theme	Domain	k/K	P value	Comment
8	crackpot theory	🔵	5/30	0.0726	Parent of Atlantis and Chariots of the Gods.
9	man vs. nature	🔴	3/13	0.0732	Nature faced: arctic, desert, volcanic morass.
10	Atlantis	🔵	1/1	0.0791	Aquan society submerged in watery cataclysm.

<div align="center">Top 20 TNG Enriched Themes</div>

Rank	Theme	Domain	k/K	P value	Comment
1	human emotion	🔴	176/265	<0.0001	Virtues, vices, emotions pleasant and unpleasant.
2	virtual reality room	🟡	31/32	<0.0001	I.e., the holodeck.
3	familial love	🔴	49/59	0.0002	Bonds between family members emphasized.
4	growing up	🔴	40/47	0.0003	Problems faced in early life treated.
5	heavenly virtue	🔴	105/144	0.0004	Emphasized human virtues detailed in main text.
6	android	🟡	33/38	0.0006	I.e., Lieutenant Commander Data.
7	human personality	🔴	176/271	0.0008	Parent theme of human emotion.

(Contd.)

Rank	Theme	Domain	k/K	P value	Comment
8	social interaction	🔴	166/249	0.0008	People interacting in groups explored.
9	pride	🔴	51/64	0.0012	Just one of the many human vices examined.
10	belonging	🔴	27/31	0.0020	Finding one's place in a group explored.
11	mother and son	🔴	24/27	0.0021	E.g. Beverly Crusher and her son Wesley.
12	introspection	🔴	106/149	0.0026	Characters keen on self-analysis of mental states.
13	maternal love	🔴	20/22	0.0029	A mother's love for her child featured.
14	AI point of view	🟢	26/30	0.0030	The world as might be viewed by an AI shown.
15	family affairs	🔴	91/126	0.0035	Ups and downs of family life examined.
16	adolescence	🔴	19/21	0.0044	Difficulties faced by teenagers explored.
17	familial relations	🔴	81/111	0.0044	Parent of mother and son/father and son.

(Contd.)

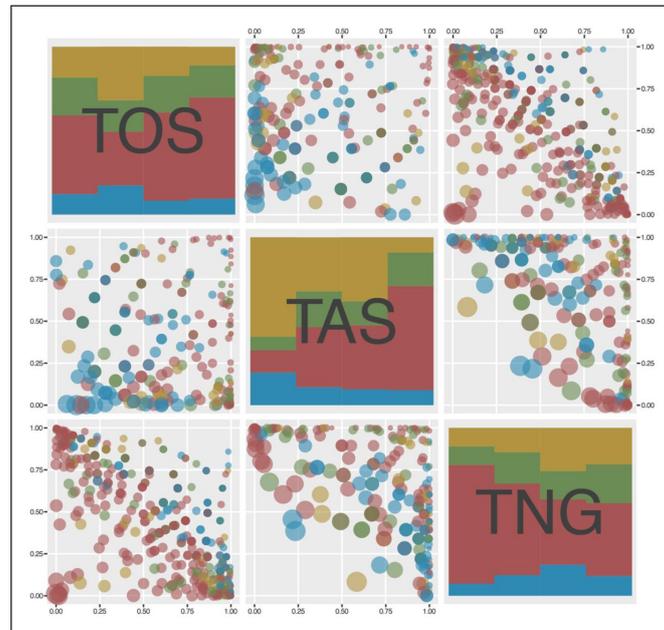| Rank	Theme	Domain	k/K	P value	Comment
18	father and son	🔴	21/24	0.0065	E.g. Worf and his son Alexander Rozhenko.
19	paternal love	🔴	23/27	0.0087	Parents' love for their children featured.
20	child rearing	🔴	10/10	0.0094	Struggles of raising a child explored.

Figure 3: Star Trek television series theme enrichment scatterplot matrix. Theme enrichment *p*-value stacked percentage bar plots for *TOS, TAS* and *TNG* are plotted along the diagonal from top left to bottom right. For each series, the *p*-values are divided into quartiles from left to right and each stacked bar indicated the percentage of themes from each domain. Theme enrichment *p*-value scores are plotted against one another for each pair of series in the off-diagonal panels. Circle size is proportional to the square root of the negative of the logarithm of the product of *p*-values pairs. Color corresponds to theme domain: the human condition (red), society (green), the pursuit of knowledge (blue), alternate reality (yellow).

of women in society (Ferguson et al. 1997; Vettel-Becker 2014). Three of the seven enriched alternate reality domain themes (i.e., "alternate society", "existential risk" and "man-made existential risk") confront viewers with ideas about how society could be changed for better or worse. In particular, the emphasis on existential risks is likely a reflection of the Cold War and relatively fresh memories of the Second World War (Cantor 2000). In light of this, it is interesting to note that five of the most enriched human condition domain themes (i.e., "wrath", "facing a fight to the death", "rage", "unpleasant emotion" and "disagreeable characteristic") are also closely tied to conflict. The remaining human condition domain themes (i.e., "way of life", "purpose in life", "personal ethical dilemma", "tough decision" and "the need for a challenge in life") pertain to life choices and decision making. Speculating about why this might be a feature of *TOS* relative to the later series is left to the reader. Suffice it to say that they, like all the most enriched human condition domain themes in *TOS*, are notably different from the most enriched themes in *TNG*.

TAS: Seven of the ten most enriched themes can be labeled simply as fanciful notions. Four are typical sci-fi themes: "earth-life inspired life form", "life-support belt", "miscellaneous life form" and "what if my life were different". The remaining two themes "Chariots of the Gods" and "Atlantis" refer to the "crackpot theories" that aliens made contact with humans in ancient times and Atlantis was a civilization with advanced technology, respectively. That such themes are enriched in *TAS* can be explained by the fact that it is the only animated series of the trio. This would have released the authors from constraints otherwise imposed by the need for costly props and special effects (see **Table 5** for some examples) and allowed them to further unleash their imagination. We hypothesize the lack of emotion-related themes may be partially explained by early animation technology's inability to approach the nuances of facial expression and body language that the consummate actors of *TOS* and *TNG*, such as William Shatner and Patrick Stewart, would routinely employ.

TNG: Nearly all of the 20 most enriched *TNG* themes relate to individual human experience. About half of them are descendants of the themes "family affairs":

"familial love", "growing up", "mother and son", "maternal love", "adolescence", "familial relations", "father and son", "paternal love" and "child rearing". The others are not dissimilar: "human emotion", "heavenly virtue", "human personality", "social interaction", "pride", "belonging" and "introspection". Anyone familiar with the android Data will recognize that the themes "android" and "AI point of view" relate to stories about individual human experience as well. One theme that stands out for not falling under the human condition is "virtual reality room", a particular sci-fi concept that refers to the holodeck in *TNG*, which has become something of a meme in its own right. Why it is that *TNG* so distinctly features these family affair, relationship, and emotional themes is of course open to interpretation. We speculatively hypothesize that an inexorable trend in modern television has been towards vapid character development designed to evoke safe familiarity rather than intellectual stimulus or moral controversy. Be that as it may, it is safe to say that the main characters in *TNG* have more elaborate background stories, subtle personality traits and complicated interpersonal relationships than the main characters of the two earlier series.

Negative control experiments

We performed a series of negative control experiments. In one such experiment, we performed enrichment analyses for 1000 test storysets consisting of n = 8 randomly selected *TOS/TAS* episodes against the same background. This mirrors the *TOS/TAS* Klingon positive control settings. The average number of significantly enriched themes was 10±5 at significance level α = 0.05. The corresponding Klingon storyset has 28 significantly enriched themes at the same significance level. This is noticeably more enriched themes on average than would be expected by chance. In a similar negative control with n = 18 for the test storyset, we found the average number of significantly enriched themes to be 13 ± 5 relative to a background of all *TNG* episodes. The corresponding Klingon storyset has 22 significantly enriched themes. Again, this is more than would be expected to be enriched by chance.

Hypergeometric test and TF-IDF comparison

TF-IDF is a statistic that is used in data mining to measure the importance of a word in a document in a collection of documents (Rajaraman and Ullman 2011). It is implemented in such textual analysis tools as the Stanford Topic Modeling Toolbox (Ramage et al. 2009), TAPoR (TAPoR 3 2019), TOME (Klein et al. 2015), Word Seer (Muralidharan et al. 2013) and Voyant Tools (Sinclair et al. 2019). In this subsection, we take a TF-IDF approach to the identification of enriched themes in the Klingon and series case studies, and compare the results with those obtained using the hypergeometric test. We implemented the TF-IDF formula $-k/n \times \log(K/N)$, where k/n is the term frequency, and $-\log(K/N)$ is a version of the logarithmically scaled inverse document frequency. **Table 6** contains the top 20 TF-IDF scoring themes in *TOS*/*TAS* Klingon and *TNG* Klingon episodes. **Table 7** contains the analogous results for the individual series *TOS*, *TAS* and *TNG*. Tables containing the full results from these analyses are included in Supplementary Information File 3 (Sheridan and Onsjö 2019B). The results reported in Tables 6 and 7 significantly overlap with those obtained using the hypergeometric test. For the Klingons of *TOS*/*TAS* a remarkable 20 out of the top 20 themes are held in common (it turns out to be 38 out of the top 50), and for the Klingons of *TNG* the figure amounts to 14 out of 20. The numbers for the individual series *TOS*, *TAS* and *TNG* are 12/20, 6/10 and 8/20, respectively. The scatterplots of **Figure 4** show broad correlation between TF-IDF scores and hypergeometric test logarithmically scaled *p*-values. What we take from this is that TF-IDF and the hypergeometric test approaches agree in their broad outlines, but differ in the details. Based on these outcomes, we contend that elucidating the mathematical relationship between TF-IDF and the hypergeometric test would make for an interesting future work. For the present, we merely point out that one advantage of the hypergeometric test is that the p-values obtained therefrom are, on the face of it, more amenable to interpretation than TF-IDF scores. To be specific, the p-value of a theme is interpretable as the probability of drawing a sample consisting of at least k stories that feature the theme in n draws, without replacement, from a background of N stories of which exactly K feature the theme. By contrast, the search

Table 6: TF-IDF scores for themes in Klingon-centric episodes relative to *TOS/TAS* and *TNG* backgrounds, respectively.

Rank	Theme	Domain	k/K	TF-IDF	Comment
	Top 20 Enriched Themes in *TOS/TAS* Klingon Episodes				
1	über-belligerent alien	🟡	5/5	1.88	The Klingons are belligerence personified.
2	diplomatic negotiating	🟢	4/7	1.34	Teetered on brink of war with the Federation.
3	culturally distinguished life form	🟡	6/20	1.22	Parent theme of über-belligerent alien.
4	man vs. beast	🔴	3/5	1.13	Beasts faced: mugato (×1), tribble (×2).
5	diplomacy	🟢	5/19	1.05	Parent theme of diplomatic negotiating.
6	pacifism	🔴	3/7	1.00	Pacifists made easy targets for conquest.
7	military tactics	🔵	3/7	1.00	Masters of the art of war.
8	conflict over a shared resource	🟢	2/2	0.98	Vied with the Federation over galactic resources.
9	atrocities of war	🟢	2/2	0.98	Not above committing war crimes.
10	tribble	🟡	2/2	0.98	Mortal enemy of the Klingons.
11	war	🟢	5/22	0.96	Reveled in warfare and conquest.
12	miscellaneous life form	🟡	3/8	0.95	Parent theme of tribble.

(Contd.)

Rank	Theme	Domain	k/K	TF-IDF	Comment
13	imperialistic society	🟢	3/8	0.95	Klingon Empire subjugated weaker neighbors.
14	the art of war	🔵	5/23	0.93	Parent theme of military tactics.
15	cross cultural understanding	🟢	4/16	0.93	Parent theme of conflict of moral codes.
16	humility	🔴	3/9	0.91	Exhibited in the face of Klingon aggression.
17	conflict of moral codes	🟢	2/3	0.88	Klingon imperialism vs. Federation benevolence.
18	patience	🔴	3/10	0.87	Exhibited in the face of Klingon aggression.
19	transnational social issue	🟢	6/32	0.87	Parent theme of diplomacy and war.
20	temperance	🔴	2/4	0.81	Counterposed with Klingon licentiousness.
Top 20 Enriched Themes in *TNG* Klingon Episodes					
1	über-belligerent alien	🟡	14/14	1.97	The Klingons are still belligerence personified.
2	honor	🔴	13/18	1.65	Central to the *TNG* Klingon way of life.
3	culturally distinguished life form	🟡	14/53	0.93	Parent theme of über-belligerent alien.
4	the need for cultural heritage	🔴	5/8	0.86	Worf cherished his Klingon cultural heritage.

<div align="right">(Contd.)</div>

Rank	Theme	Domain	k/K	TF-IDF	Comment
5	rage	🔴	7/23	0.79	Flying into a violent rage is a key *TNG* Klingon trait.
6	belonging	🔴	7/27	0.73	Worf was the lone Klingon in Starfleet.
7	father and son	🔴	6/21	0.71	E.g. Worf and his son Alexander.
8	the art of war	🔵	6/24	0.66	Still masters of the art of war.
9	loyalty	🔴	7/33	0.65	A ruthless yet loyal alien race.
10	surprise	🔴	6/27	0.62	E.g. Worf was surprised to find he had a young son.
11	wrath	🔴	8/44	0.62	Flying into a violent rage is a key *TNG* Klingon trait.
12	guilt and evidence	🟢	4/11	0.62	Applies by and large to non-Klingon characters.
13	cooperation	🔴	5/20	0.60	Klingon Empire and Federation sometimes joined forces.
14	diligence	🔴	7/38	0.60	Mostly unrelated to Klingon activity.
15	racism in society	🟢	3/5	0.59	Worf despised the Romulan race.
16	facing wrongful accusations	🔴	3/5	0.59	E.g. Worf falsely accused of treason.
17	human need	🔴	9/54	0.59	Parent theme of honor.

(Contd.)

Rank	Theme	Domain	k/K	TF-IDF	Comment
18	disgust	🔴	5/22	0.58	Mostly unrelated to Klingon activity.
19	personal conviction	🔴	5/23	0.57	E.g. a trio of Klingon fanatic stayed true to their holy cause.
20	the lust for power	🔴	3/7	0.54	Various Klingons lusted to rule Klingon Empire.

Table 7: TF-IDF scores for themes in each Star Trek television series relative to a *TOS/TAS/TNG* background.

Rank	Theme	Domain	k/K	TF-IDF	Comment
Top 20 *TOS* Enriched Themes					
1	wrath	🔴	41/85	0.61	Violent outbursts abounded.
2	gender issues	🟢	26/46	0.58	Challenged some outmoded sexist attitudes.
3	female stereotype	🟢	19/24	0.58	Reinforced some outmoded female stereotypes, too.
4	rage	🔴	27/50	0.58	Captain Kirk no stranger to such altercations.
5	alternate society	🟡	45/100	0.58	Mind-openingly different societies explored.
6	alternate life	🟡	39/87	0.57	Fantastical things befalling people explored.
7	existential risk	🟡	37/83	0.56	Various threats to human civilization explored.
8	what if I faced an incredible situation	🟡	30/64	0.55	Examples of incredible situations faced are provided in the main text.

(Contd.)

Rank	Theme	Domain	k/K	TF-IDF	Comment
9	way of life	🔴	24/45	0.55	Ideas on how to lead a good life explored.
10	past and present moral controversies	🟢	24/49	0.52	Social issues of the day explored.
11	man-made existential risk	🟡	22/43	0.51	E.g. the dangers of WMDs and societal laziness.
12	ethics	🔵	31/74	0.51	Confronts viewers with various ethical questions.
13	personal ethical dilemma	🔴	45/112	0.51	Characters faced with difficult moral choices.
14	serenity	🔴	21/41	0.50	E.g. the Captain Kirk self-hug scene in *TOS*3x03.
15	disagreeable characteristic	🔴	48/121	0.50	E.g. complacency, deviousness, rudeness, etc.
16	social order	🟢	22/48	0.48	Various ways of organizing society explored.
17	lust	🔴	28/70	0.48	E.g. Captain Kirk's evil half lusted after Janice Rand in *TOS*1x04.
18	compassion	🔴	25/60	0.48	E.g. Captain Kirk's good half oozed with compassion in *TOS*1x04.
19	personal practical dilemma	🔴	26/64	0.48	Individual tough decisions faced.
20	amicable disposition	🔴	45/119	0.48	E.g. magnanimity in victory and graciousness in defeat.
	Top 10 *TAS* Enriched Themes				
1	earth-life inspired life form	🟡	6/15	0.80	Avians, felinoids, insectoids, slug-like aliens, etc.

Rank	Theme	Domain	k/K	TF-IDF	Comment
2	miscellaneous life form	◯	5/13	0.70	E.g. an alien composed of autonomous parts.
3	life-support belt	◯	3/3	0.62	A belt-like device that functions as a spacesuit.
4	what if my life were different	◯	6/38	0.54	E.g. what if I were a specimen animal in a zoo.
5	alternate society	◯	11/100	0.51	E.g. a master race controlled society.
6	crackpot theory	●	5/30	0.51	E.g. aliens supplied the ancients with technology.
7	existential risk	◯	9/83	0.49	E.g. Aquan society submerged in watery cataclysm.
8	man vs. beast	●	3/8	0.48	Beasts faced: ie-matya, rock beast, tribble.
9	alternate life	◯	9/87	0.48	Parent theme of what if my life were different.
10	what's out there	◯	6/51	0.46	E.g. an asteroid belt, a supernova and a space aurora.
Top 20 *TNG* Enriched Themes					
1	familial love	●	49/59	0.43	Bonds between family members emphasized.
2	pride	●	51/64	0.43	Just one of the many human vices examined.
3	familial relations	●	81/111	0.42	Virtues, vices, emotions pleasant and unpleasant.
4	family affairs	●	91/126	0.41	Ups and downs of family life examined.
5	growing up	●	40/47	0.40	Problems faced in early life treated.

(Contd.)

Rank	Theme	Domain	k/K	TF-IDF	Comment
6	culturally distinguished life form	🟡	53/73	0.40	E.g. xenophobic isolationist alien races.
7	group interaction	🔴	66/96	0.40	Parent theme of family affairs.
8	heavenly virtue	🔴	105/144	0.39	Emphasized human virtues detailed in main text.
9	human nature	🔵	57/83	0.39	The human condition explored.
10	AI technology	🟡	45/61	0.39	Parent theme of android.
11	state of mind	🔴	81/120	0.39	Characters' mental well-being routinely explored.
12	views and customs	🟢	88/129	0.38	*TNG* embraced multiculturalism.
13	fictional apparatus	🟡	54/80	0.38	Parent theme of virtual reality room.
14	annoyance	🔴	56/84	0.38	Just one of the many human emotions explored.
15	virtual reality room	🟡	31/32	0.38	I.e., the holodeck.
16	introspection	🔴	106/149	0.38	Characters keen on self-analysis of mental states.
17	android	🟡	33/38	0.37	I.e., Lieutenant Commander Data.
18	coping with an affliction	🔴	46/67	0.37	Sickbay a carousel for critically ill people.
19	agreeable characteristic	🔴	96/142	0.37	E.g. courage.
20	intercultural relations	🟢	95/141	0.37	*TNG* embraced multiculturalism.
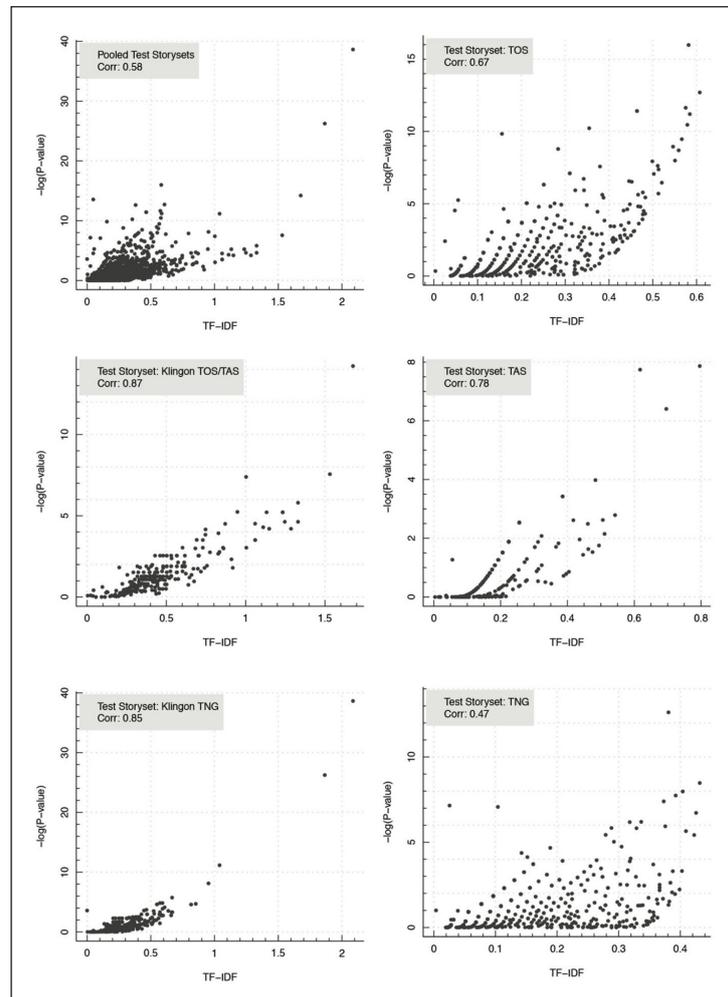
Figure 4: Scatterplots of TF-IDF scores versus hypergeometric test negative logarithm *p*-values for themes from each enrichment analysis from Sections 5.1 and 5.2 of the main text. The upper left scatterplot shows the pooled results from all five cases.

for a simple probabilistic interpretation of TF-IDF scores remains a subject of ongoing research (Spärck Jones 1972; Robertson 2004; Havrlant and Kreinovich 2017).

Discussion

The primary aim of this paper has been to introduce the hypergoemetric test for theme enrichment analysis to the digital literary studies community. We consider our proposed draft theme ontology and toy Star Trek thematic dataset to be

contributions of a secondary nature. The hypergeometric testing approach to the identification of enriched themes in a list of stories equips the literary scholar with a new weapon to wield in their fiction document analyses. We have demonstrated the potential of the hypergeometric test by applying it to Star Trek television franchise thematic data. In the first place, we found the lists of enriched themes produced by the test to be helpful for identifying what makes the corresponding storysets special and for generating speculative hypotheses. In the second place, we argued that the hypergeometric test compares favorably with the commonly used TF-IDF statistic when it comes to answering the question of whether such-and-such a theme (or any term for that matter) occurs in a storyset at a frequency significantly greater than would be expected by chance. But we would be remiss not to touch on the potential for "garbage in, garbage out" (GIGO) to bias and confound theme enrichment analyses. In other words, the result of a theme enrichment analysis using the hypergeometric test is at best as good as the themes on which it is based. The GIGO menace warrants special mention in the context of thematic analyses. The reason is that themes, unlike objective scientific entities such as genes, are generally complex abstractions of a highly subjective nature. We, therefore, stress the preliminary nature of our work in applying the hypergeometric test to thematic data. All that said, it is our hope that this work will contribute to the judicious use of the hypergeometric test in the digital literary studies in the fullness of time.

Two main obstacles stand in the way of making our approach to theme enrichment analysis practical on a large-scale. First, a protocol for annotating stories with themes in such a manner that stories can be meaningfully compared in terms of their shared themes must be developed. We have taken a first step toward addressing this need by proposing a Basic Formal Ontology compliant draft theme ontology. Moving forward, we aim to integrate related ontologies such as the Emotion Ontology (Hastings et al. 2011) to name but one. Ontology design is an open-ended process, subject to setbacks and changes of direction. It is plain that our draft theme ontology will be no exception. However, we point out that even if the structure of the ontology changes markedly, many of the themes will remain

intact as presently defined. Second, a large-scale database of compatibly themed stories is required. To this end, we have launched the Theme Ontology (beta version) online community platform (Theme Ontology 2019). The website features an ever-expanding controlled vocabulary of defined themes, hierarchically arranged into our draft theme ontology. Community members are encouraged to tag whatever stories (e.g., short stories, novels, films, TV shows, etc.) they please with themes drawn from the ontology, and adorn the ontology with newly coined themes as necessary. Stories are manually tagged with themes at present. Topic modeling techniques (Blei 2012) as implemented in such software packages as MALLET (McCallum 2002), the Python module gensim (Řhůřek and Sojka 2010), topicmodels (Grün and Hornik 2011) and the R package tm (Feinerer et al. 2008) have been used successfully to identify literary themes in text copora (Jockers 2013; Jockers and Mimno 2013; Goldstone and Underwood 2014; Boyd-Graber et al. 2017). In the future, we plan to use topic modeling to automatically collect themes for large numbers of stories in order to grow the Theme Ontology database. An interesting challenge awaits in figuring out how to adapt the current methods for automatic topic labeling (Lau et al. 2011; Cano Basave et al. 2014; Bhatia et al. 2017) to the problem of mapping identified topics to LTO themes. Lastly, a theme enrichment analyzer web application is available for download at the Theme Ontology GitHub repository (Theme Ontology Project GitHub Repository 2019). Tools from the stoRy package, including our theme enrichment test, will be made accessible as web applications on the Theme Ontology website in order to help users analyze curated thematic datasets. It is our aim to build up a large-scale database of freely available thematically annotated stories that can be analyzed using web applications within the Theme Ontology ecosystem.

Concerning the identification of themes in stories, we emphasize that our annotating of Star Trek television series episodes with themes was manual, subjective to our point of view, and not fully replicable. In general, theme identification is admittedly subjective, but this is not to say the endeavor is altogether arbitrary. The Lexico dictionary definition of theme quoted in the introduction, which proves adequate for most literary-critical purposes, is problematic from the present point of view insofar as it designates a theme to be a property of a story. Instead, we propose

to consider a theme as a relation between story and partaker thereof. According to this subjective view, it is entirely possible for partaker A to contend that theme X is featured in story Y, but not partaker B. Objectivity is approached to the extent that universal agreement among story partakers is attained. The potential for a system of this kind to degenerate into a wasteland of subjectivity depends on the extent to which themes can be made precise. At the Theme Ontology community platform, we are presently drafting a policies and guidelines document that will emphasize the need for clarity in and verifiability of theme definitions. The theme "the desire for vengeance", which is defined as "A character seeks retribution over a perceived injury or wrong.", constitutes a model definition. Growing pains are inevitable. But by concentrating on cataloging precisely defined and verifiable themes, i.e., the low-hanging fruit, we hope to ensure that LTO becomes a useful literary studies resource in the future.

There are a few points to mention in closing. First, the toy thematic dataset we have introduced in the present work consists of the combined 280 episodes of *TOS*, *TAS* and *TNG*. We contend this sufficed for our purpose of demonstrating the value of theme enrichment analysis. But the entire Star Trek franchise is made up of 754 episodes and 14 films. A good future work would be to round out the toy dataset with themed episodes from all the other series along with the films. Second, the hypergeometric test is designed to answer the question: what themes in a test storyset of interest stand out against a background storyset. It would be interesting to extend the test to so as identify enriched co-occurrences of themes in stories. Other literary questions are also amenable to statistical investigation. For example, the investigator may wish to discover a subset of stories in a storyset that have similar themes by performing a clustering analysis. Another example is time-series analysis for the study of how theme usage changes over time in a storyset with timestamped stories. The stoRy package should be extended to include statistical methods to address questions of these sorts. Finally, we note that we have recently advanced an LTO theme exploiting story recommender system (Sheridan, Onsjö, and Hastings 2019). Third, the version of LTO used in this work features about 1500 unique themes, but smaller collections of themes may prove more suitable for certain applications. The

same issue arises in genetic testing where only small subsets out of roughly twenty thousand genes need to be examined at one time. Gene panels, which are curated sets of genes related to particular conditions, are used in genetic testing to overcome this problem (Hall et al. 2014). In the future, we plan to curate "theme panels", by which we mean sets of themes related to specific topics (e.g., "artificial intelligence", "crime and punishment", and "family life"). In this manner, scholars will be able to concentrate on those themes relating to their specific research hypotheses. Lastly, computer-assisted qualitative data analysis software is sometimes used for document analysis in the social sciences. ATLAS.ti (Muhr 2004), NVivo (QSR International Pty Ltd. 2015) and MAXQDA (VERBI Software 1989-2019) are three such programs that allow users to manage, analyze and visualize data related to text, audio, and video documents. It is possible to use these sorts of programs to annotate stories with themes and subsequently explore the correlations among them. However, to our knowledge, none of these programs implement the hypergeometric test. The same thing holds for the popular textual analysis tools the Stanford Topic Modeling Toolbox (Ramage et al. 2009), TAPoR (TAPoR 3 2019), TOME (Klein et al. 2015), Word Seer (Muralidharan et al. 2013) and Voyant Tools (Sinclair et al. 2019). It would be profitable to augment these programs with a hypergeometric test function for term enrichment.

Additional Files

The additional files for this article can be found as follows:

- **List of Klingon related Star Trek episodes.** A list of Star Trek TOS/TAS/TNG television series episodes featuring the Klingon alien race. The criterion for inclusion is that the Klingons were deemed by the authors to have been featured throughout the episode in a way that is central to the story plot. DOI: https://doi.org/10.16995/dscn.316.s1
- **Hypergeometric test results.** Tables of over-represented Literary Theme Ontology version 0.1.1 literary themes in Star Trek TOS/TAS/TNG television series storysets as identified by the hypergeometric test. DOI: https://doi.org/10.16995/dscn.316.s2

- **TF-IDF results.** Tables of over-represented Literary Theme Ontology version 0.1.1 literary themes in Star Trek TOS/TAS/TNG television series storysets as identified by the TF-IDF statistic. DOI: https://doi.org/10.16995/dscn.316.s3
- **Supplementary Material.** Frontiers. DOI: https://doi.org/10.16995/dscn.316.s4

Acknowledgements

Competing interests

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Author contributions

MO and PS collected the data. MO and PS conceived the analysis. PS conducted the analysis. MO and PS wrote the manuscript.

Editorial contributions

Section/Copy editor: Nathir Haimoun, University of Lethbridge Journal Incubator.

Copy editor: Shahina Parvin, University of Lethbridge Journal Incubator.

Works cited

Arp, Robert, Barry Smith, and **Andrew D. Spear.** 2015. *Building Ontologies with Basic Formal Ontology*. London: The MIT Press. DOI: https://doi.org/10.7551/mitpress/9780262527811.001.0001

BBC (British Broadcasting Corporation). 2011. "Klingons and Commies." *BBC Online*. Accessed November 10, 2019. http://bbc.adactio.com/cult/st/original/commies.shtml.

Benjamini, Yoav, and **Daniel Yekutieli.** 2001. "The Control of the False Discovery Rate in Multiple Testing under Dependency." *Annals of Statistics* 29(4): 1165–1188. DOI: https://doi.org/10.1214/aos/1013699998

Benjamini, Yoav, and **Yosef Hochberg.** 1995. "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing." *Journal of the Royal Statistical Society. Series B (Methodological)* 57(1): 289–300. DOI: https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Bhatia, Shraey, Jey Han Lau, and **Timothy Baldwin.** 2017. "An Automatic Approach for Document-level Topic Model Evaluation." *Proceedings of the 21st Conference on Computational Natural Language Learning (CoNLL 2017)*, Vancouver, Canada, August 3–4, 206–215. Vancouver : Association for Computational Linguistics. DOI: https://doi.org/10.18653/v1/K17-1022

Blair, Karin. 1983. "Sex and 'Star Trek' (Le sexe dans 'Star Trek')." *Science Fiction Studies* 10(3): 292–297.

Blei, David M. 2012. "Probabilistic Topic Models." *Communications of the ACM* 55 (4): 77–84. DOI: https://doi.org/10.1145/2133806.2133826

Bondurant, Tom. 2017. "From Kor to Discovery, We Say Qapla' To 50 Years of Klingons." *CBR.* Accessed November 10, 2019. https://www.cbr.com/star-trek-klingons-50-anniversary.

Boyd-Graber, Jordan, Yuening Hu, and **David Mimno.** 2017. "Applications of Topic Models." *Foundations and Trends in Information Retrieval* 11(2–3): 143–296. DOI: https://doi.org/10.1561/1500000030

Boyle, Elizabeth I., Shuai Weng, Jeremy Gollub, Heng Jin, David Botstein, J. Michael Cherry, and **Gavin Sherlock.** 2004. "GO::TermFinder--Open Source Software for Accessing Gene Ontology Information and Finding Significantly Enriched Gene Ontology Terms Associated with a List of Genes." *Bioinformatics* 20(18): 3710–3715. DOI: https://doi.org/10.1093/bioinformatics/bth456

Cano Basave, Amparo Elizabeth, Yulan He, and **Ruifeng Xu.** 2014. "Automatic Labelling of Topic Models Learned from Twitter by Summarisation." In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics, ACL 2014* (Volume 2: Short Papers), Baltimore, USA, June 22–27, 618–624. Baltimore: Association for Computational Linguistics. DOI: https://doi.org/10.3115/v1/P14-2101

Cantor, Paul A. 2000. "Shakespeare in the Original Klingon: Star Trek and the End of History." *Perspectives on Political Science* 29(3): 158–166. DOI: https://doi.org/10.1080/10457090009600707

Comprehensive R Archive Network, The. 2019. "The Comprehensive R Archive Network." Accessed November 15. https://cran.r-project.org/.

Feinerer, Ingo, Kurt Hornik, and **David Meyer.** 2008. "Text Mining Infrastructure in R." *Journal of Statistical Software* 25(5): 1–54. DOI: https://doi.org/10.18637/jss.v025.i05

Ferguson, Kathy E., Gilad Ashkenazi, and **Wendy Schultz.** 1997. "Gender Identity in Star Trek." In *Political Science Fiction,* edited by Donald M. Hassler, and Clyde Wilcox, 214–233. Columbia, SC: University of South Carolina Press.

GitHub. 2019. "Theme Ontology Project GitHub Repository." Accessed November 15. https://github.com/theme-ontology/shiny-apps.

Goldstone, Andrew, and **Ted Underwood.** 2014. "The Quiet Transformations of Literary Studies: What Thirteen Thousand Scholars Could Tell Us." *New Literary History* 45(3): 359–384. DOI: https://doi.org/10.1353/nlh.2014.0025

Gonzalez, George A. 2015. *The Politics of Star Trek: Justice, War, and the Future.* New York: Palgrave Macmillan. DOI: https://doi.org/10.1057/9781137546326

Griffith, Kelley. 2010. *Writing Essays about Literature.* 8th ed. Belmont, CA: Wadsworth Publishing.

Grün, Bettina, and **Kurt Hornik.** 2011. "topicmodels: An R Package for Fitting Topic Models." *Journal of Statistical Software* 40(13): 1–30. DOI: https://doi.org/10.18637/jss.v040.i13

Hall, Michael J., Andrea D. Forman, Robert Pilarski, Georgia Wiesner, and **Veda N. Giri.** 2014. "Gene Panel Testing for Inherited Cancer Risk." *Journal of the National Comprehensive Cancer Network* 12(9): 1339–1345. DOI: https://doi.org/10.6004/jnccn.2014.0128

Hastings, Janna, Werner Ceusters, Barry Smith, and **Kevin Mulligan.** 2011. "The Emotion Ontology: Enabling Interdisciplinary Research in the Affective Sciences." In *Proceedings of the 7th International and Interdisciplinary Conference on Modeling and Using Context (CONTEXT 2011),* Karlsruhe, Germany,

September 26–30, 119–123, Berlin, Heidelberg: Springer-Verlag. DOI: https://doi.org/10.1007/978-3-642-24279-3_14

Havrlant, Lukáš, and **Vladik Kreinovich.** 2017. "A Simple Probabilistic Explanation of Term Frequency-Inverse Document Frequency (tf-idf) Heuristic (and Variations Motivated by this Explanation)." *International Journal of General Systems* 46(1): 27–36. DOI: https://doi.org/10.1080/03081079.2017.1291635

Haynes, Natalie. 2016. "Star Trek and the Kiss That Changed TV." *BBC.* Accessed November 10, 2019. http://www.bbc.com/culture/story/20160707-star-trek-turns-50-why-it-was-subversive-and-groundbreaking.

Hitzler, Pascal, Markus Krötzsch, Bijan Parsia, Peter F. Patel-Schneider, and **Sebastian Rudolph** (eds.). 2012. *OWL 2 Web Ontology Language Primer.* W3C Recommendation. 2nd ed. Accessed November 10, 2019. https://www.w3.org/TR/owl2-primer/.

Huang, Da Wei, Brad T. Sherman, and **Richard A. Lempicki.** 2009. "Bioinformatics Enrichment Tools: Paths Toward the Comprehensive Functional Analysis of Large Gene Lists." *Nucleic Acids Research* 37(1): 1–13. DOI: https://doi.org/10.1093/nar/gkn923

Hudson, William Henry. 1913. *An Introduction to the Study of Literature.* 2nd ed. London, UK: George G. Harrap & Company.

Jockers, Matthew L. 2013. *Macroanalysis: Digital Methods and Literary History.* Ist ed. Champaign, IL, USA: University of Illinois Press. 1st edn. DOI: https://doi.org/10.5406/illinois/9780252037528.001.0001

Jockers, Matthew L., and **David Mimno.** 2013. "Significant Themes in 19th-century Literature." *Poetics* 41(6): 750–769. DOI: https://doi.org/10.1016/j.poetic.2013.08.005

Klein, Lauren F., Jacob Eisenstein, and **Iris Sun.** 2015. "Exploratory Thematic Analysis for Digitized Archival Collections." *Digital Scholarship in the Humanities* 30(suppl_1): i130–i141. DOI: https://doi.org/10.1093/llc/fqv052

Knight, Graham, and **Jennifer Smith.** 1998. "High-tech Feudalism: Warrior Culture and Science Fiction TV." *Florilegium.*15: 267–282.

Lau, Jey Han, Karl Grieser, David Newman, and **Timothy Baldwin.** 2011. "Automatic Labelling of Topic Models." In *The 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies, Proceedings of the Conference*, Portland, Oregon, USA, 19–24 June, 1536–1545. Stroudsburg, PA: The Association for Computational Linguistics.

Lexico. 2019. Accessed November 10. https://www.lexico.com/en/definition/knowledge.

McCallum, Andrew K. 2002. "MALLET: Machine Learning for Language Toolkit." *MALLET*. Accessed November 10, 2019. http://mallet.cs.umass.edu/.

McDonald, John H. 2014. *Handbook of Biological Statistics*. 3rd ed. Baltimore: Sparky House Publishing.

Meijer, Rosa J., and **Jelle J. Goeman.** 2016. "Multiple Testing of Gene Sets from Gene Ontology: Possibilities and Pitfalls." *Briefings in Bioinformatics* 17(5): 808–818. DOI: https://doi.org/10.1093/bib/bbv091

Muhr, Thomas. 2004. "User's Manual for ATLAS.ti 5.0." Berlin: ATLAS.ti Scientific Software Development GmbH.

Muralidharan, Aditi, Marti A. Hearst, and **Christopher Fan.** 2013. "Wordseer: A Knowledge Synthesis Environment for Textual Data." In *Proceedings of the 22nd ACM International Conference on Information and Knowledge Management (CIKM'13)*, San Francisco, CA, USA, October 27–November 1, 2533–2536. New York, NY: Association for Computing Machinery. DOI: https://doi.org/10.1145/2505515.2508212

Noble, William S. 2009. "How does Multiple Testing Correction Work?" *Nature Biotechnology* 27(12): 1135–1137. DOI: https://doi.org/10.1038/nbt1209-1135

NVivo Blog, The. 2019. "What Is Qualitative Research?" Accessed November, 12. http://www.qsrinternational.com/nvivo/nvivo-community/blog/what-is-qualitative-research.

QSR International Pty Ltd. 2015. NVivo Qualitative Data Analysis Software Version 11. Accessed November 12, 2019. https://www.qsrinternational.com/nvivo/home.

Rajaraman, Anand, and **Jeffrey David Ullman.** 2011. *Mining of Massive Datasets.* New York: Cambridge University Press.

Ramage, Daniel, Evan Rosen, Jason Chuang, Christopher D. Manning, and **Daniel A. McFarland.** 2009. *Topic Modeling for the Social Sciences.* In *NIPS 2009 Workshop on Applications for Topic Models: Text and Beyond* 5: 1–4. Accessed November 11, 2019. https://nlp.stanford.edu/dramage/papers/tmt-nips09.pdf.

Řhůřek, Radim, and **Petr Sojka.** 2010. "Software Framework for Topic Modelling with Large Corpora." In *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*, Valletta, Malta, May 22, 46–50. Accessed November 15, 2019. http://www.lrec-conf.org/proceedings/lrec2010/workshops/W10.pdf

Robertson, Stephen. 2004. "Understanding Inverse Document Frequency: On Theoretical Arguments for IDF." *Journal of Documentation* 60(5): 503–520. DOI: https://doi.org/10.1108/00220410410560582

Roblin, Sebastien. 2016. "Star Trek's Original Series Brought the Cold War into Space." *Medium.* October 28. Accessed November 15, 2019. https://medium.com/war-is-boring/star-treks-original-series-brought-the-cold-war-into-space-91ae5a9291f4.

Salton, Gerard, A. Wong, and **Chung-Shu Yang.** 1975. "A Vector Space Model for Automatic Indexing." *Communications of the ACM* 18(11): 613–620. DOI: https://doi.org/10.1145/361219.361220

Science Fiction Encyclopedia, The (SFE). 2019. Accessed November, 12. http://www.sf-encyclopedia.com/.

Sheridan, Paul, and **Mikael Onsjö.** 2017. "stoRy: Functions for the Analysis of Star Trek Thematic Data. R package version 0.1.1." Accessed November 11, 2019. https://CRAN.R-project.org/package=stoRy.

———. 2019A. theme-ontology/theming: LTO 0.1.1 [Dataset]. *Zenodo.* DOI: http://doi.org/10.5281/zenodo.3522499

———. 2019B. "Supplementary Tables for 'Theme Enrichment Analysis: A Statistical Test for Identifying Significantly Enriched Themes in a List of Stories with an Application to the Star Trek Television Franchise'" [Dataset]. *Zenodo.* DOI: https://doi.org/10.5281/zenodo.3357174

Sheridan, Paul, Mikael Onsjö, Claudia Becerra, Sergio Jimenez, and **George Dueñas.** 2019. "An Ontology-Based Recommender System with an Application to the Star Trek Television Franchis*e*." *Future Internet* 11(9): 182. DOI: https://doi.org/10.3390/fi11090182

Sheridan, Paul, Mikael Onsjö, and **Janna Hastings.** 2019. "The Literary Theme Ontology for Media Annotation and Information Retrieval." *First International Workshop on Ontologies for Digital Humanities and their Social Analysis*, Graz, Austria, September 23–15.

Sinclair, Stefan, Geoffrey Rockwell, and **the Voyant Tools Team.** 2019. "Voyant: See Through Your Text." *Voyant Tools.* Accessed November 15. https://voyant-tools.org.

Smith, Barry, Anand Kumar, and **Thomas Bittner.** 2005. "Basic Formal Ontology for bioinformatics: IFOMIS Reports."

Spärck Jones, Karen. 1972. "A Statistical Interpretation of Term Specificity and its Application in Retrieval." *Document Retrieval Systems* 28(1): 11–21. DOI: https://doi.org/10.1108/eb026526

STARFLEET. 2019. STARFLEET: the International Star Trek Fan Association, Inc. Accessed November 15. http://sfi.org/.

Star Trek. 2019. CBS Studios Inc. Accessed November 15. https://www.startrek.com/.

TAPoR 3. 2019. "Discover Research Tools for Studying Texts." *TAPoR.* Accessed November 12. http://tapor.ca/tools.

Taylor, Mary P. 2002. *Adventures in Time and Space. Star Trek.* New York: Pocket Books.

Telotte, J. P. 2008. *The Essential Science Fiction Television Reader.* Essential Readers in Contemporary Media and Culture. Kentucky: University Press of Kentucky.

Theme Ontology. 2019. "Theme Ontology." Accessed November 12. https://www.themeontology.org/.

TOS (Star Trek: The Original Series). 1969. "The Devil in the Dark." Season 1, episode 25, directed by Joseph Pevney, aired September 6, on NBC.

VERBI Software. 1989–2019. "MAXQDA, software for qualitative data analysis." Accessed December 3. https://www.maxqda.com/about.

Vettel-Becker, Patricia. 2014. "Space and the Single Girl: Star Trek, Aesthetics, and 1960s Femininity." *Frontiers: A Journal of Women Studies* 35: 143–178. DOI: https://doi.org/10.5250/fronjwomestud.35.2.0143

Wikipedia. 2019. "Human Condition." Accessed November 10. https://en.wikipedia. org/wiki/Human_condition.

Zheng, Qi, and **Xiu-Jie Wang.** 2008. "GOEAST: A Web-based Software Toolkit for Gene Ontology Enrichment Analysis." *Nucleic Acids Research* 36: W358–W363. DOI: https://doi.org/10.1093/nar/gkn276