



## Using NLP to Generate MARC Summary Fields for Notre Dame's Catholic Pamphlets

Jeremiah Flannery

### Abstract:

Three NLP (Natural Language Processing) automated summarization techniques were tested on a special collection of Catholic Pamphlets acquired by Hesburgh Libraries. The automated summaries were generated after feeding the pamphlets as .pdf files into an OCR pipeline. Extensive data cleaning and text preprocessing were necessary before the computer summarization algorithms could be launched. Using the standard ROUGE F1 scoring technique, the Bert Extractive Summarizer technique had the best summarization score. It most closely matched the human reference summaries. The BERT Extractive technique yielded an average Rouge F1 score of 0.239. The Gensim python package implementation of TextRank scored at .151. A hand-implemented TextRank algorithm created summaries that scored at 0.144. This article covers the implementation of automated pipelines to read PDF text, the strengths and weakness of automated summarization techniques, and what the successes and failures of these summaries mean for their potential to be used in Hesburgh Libraries.

To cite this article:

Flannery, J. (2020). Using NLP to Generate MARC Summary Fields for Notre Dame's Catholic Pamphlets. *International Journal of Librarianship*, 5(1), 20-35.  
<https://doi.org/10.23974/ijol.2020.vol5.1.158>

To submit your article to this journal:

Go to <https://ojs.calaijol.org/index.php/ijol/about/submissions/>

# Using NLP to Generate MARC Summary Fields for Notre Dame's Catholic Pamphlets

Jeremiah Flannery

University of Notre Dame, Indiana, USA

## ABSTRACT

Three NLP (Natural Language Processing) automated summarization techniques were tested on a special collection of Catholic Pamphlets acquired by Hesburgh Libraries. The automated summaries were generated after feeding the pamphlets as .pdf files into an OCR pipeline. Extensive data cleaning and text preprocessing were necessary before the computer summarization algorithms could be launched. Using the standard ROUGE F1 scoring technique, the Bert Extractive Summarizer technique had the best summarization score. It most closely matched the human reference summaries. The BERT Extractive technique yielded an average Rouge F1 score of 0.239. The Gensim python package implementation of TextRank scored at .151. A hand-implemented TextRank algorithm created summaries that scored at 0.144. This article covers the implementation of automated pipelines to read PDF text, the strengths and weakness of automated summarization techniques, and what the successes and failures of these summaries mean for their potential to be used in Hesburgh Libraries.

**Keywords:** Natural Language Processing, Special Collections, Summarization, Extractive Summarization, Machine Learning, Catholic Pamphlet

## INTRODUCTION

Patrons want summaries. Of all the information fields available in the MARC catalog, the summary field ranks near the top. The summary field is behind only the author and title fields when it comes to importance to patrons (Lundgren and Simpson 1999). Patrons of differing backgrounds (Chinese and American patrons, faculty and graduate student patrons) all see value in bibliographic summaries (Luk 1996; Lundgren and Simpson 1997).

Summaries aid discoverability. Specifically, summaries improve indexing and search capabilities of a text (Borko and Bernier, 1975). Internal analysis (Unpublished Data, Kasten & Flannery, 2020) of English monographs showed that Notre Dame patrons check out records that include a summary field at higher frequencies, even when adjusted for estimated popularity of the monograph. Unfortunately, when special collections are brought to the library catalog, there often isn't the possibility to leverage existing summaries. In 2019, when the University of Notre Dame brought in a special collection of over 5500 Catholic Pamphlets<sup>1</sup>, our department did not expend hundreds of hours (or more!) of staff time to read, let alone write summaries of the 5500 texts. Instead, Natural Language Processing (NLP) Summarization methods that utilize machine

learning were used to generate summaries computationally.

### **Brief Introduction of NLP Summarization**

To see the ubiquity of NLP summarization today, one only needs to browse Twitter or Reddit, where bots regularly summarize Wikipedia and CNN articles and feature those summaries to readers. NLP summarization has existed since the 1950s, when basic term frequency was used to pinpoint keywords and important sentences (Luhn, 1958). For years, different methods struggled to significantly outperform even blunt summarization tricks, such as merely selecting the opening sentences as a summary for a whole article (Chiang, 2019). In the mid-90s, the Trainable Document Extractor and the ANES Text Extraction System represented significant progress for statistical text extraction. In the late 90s, large quantities of Internet information increased the desire for text summarization (Gupta & Lehal, 2010). Recently, advancements with neural network algorithms have allowed researchers to pull further ahead of past statistical methods (Aries et al., 2019). Current methods fall into two camps: abstractive and extractive.

Extractive summarization takes central sections of text, generally sentences, and stitches them together to make a whole summary. For the first half century of NLP summarization, the selection of text was guided by basic statistical methods applied to single words, phrases, and distance between word co-occurrences. Recently, extractive NLP methods have made advancements using deep learning techniques (Dong et al. 2018). While deep learning is not associated with extractive summaries the majority of the time, it is with abstractive summaries.

Abstractive summarization employs neural networks to generate novel text, instead of merely selecting text from the source document. Effective models are scaled to train on large amounts of data in order to accomplish this task, often leveraging the NLP concept of self-attention to perform the heavy lifting of the algorithm (Rush et al. 2015; Parikh et al., 2016). This project only focused on extractive summarization methods.

## **METHODS**

### **Data Collection**

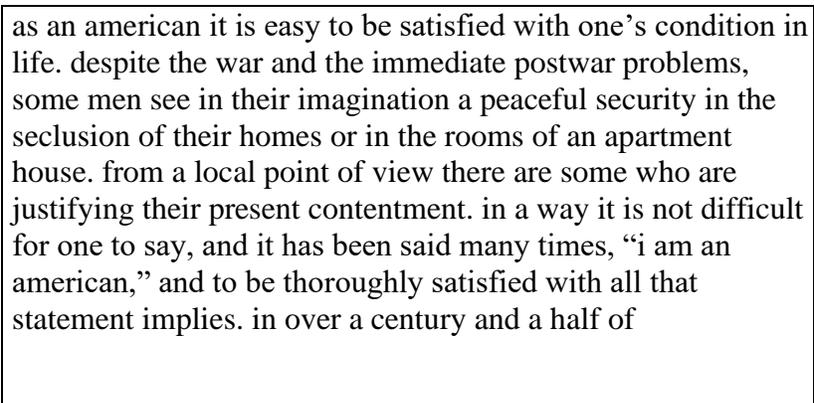
PDFs of Notre Dame's Catholic Pamphlets collection were provided by Hesburgh Library's CurateND Support Team. Existing summaries of the PDFs in this collection were sparse. While half of the collection had a MARC summary field, most of these summaries were a few words at most. A summary field containing just the two words "Catholic Pamphlet" was common. Because of this, creation of a more robust summary was attempted by combining the summaries with Library of Congress subject headings. Pamphlets with unhelpful summary fields were removed from the dataset. Fields that only contained values like "given in St Paul, MI, Jan 10 1950", "includes index", "caption: \_\_\_", "edited by \_\_\_", or simply "church" were removed as well. Occasionally, a summary field was well written and useful. The following is an example of one of the above-average summaries: "A collection of articles by John A. Elbert, Paul O'Brien, and Elmer Lackner that appeared originally in the Marianist magazine from September through December, 1949." Yet, the majority of what remained only had note fields and subject headings, for example: "Bible, Criticism, interpretation, etc., Catholic authors. Catholic Church, Apologetic works."

Of the 5500 texts, only 51 came with sufficient summaries. Another 50 summaries were written from scratch during this project, making sure to use language present in the texts when summarizing their main ideas. Computer-generated summaries of the pamphlets were compared against human-written summaries on this subset of 101 pamphlets (Appendix A).

### Text Extraction and Cleaning

The 5500 pamphlets were originally provided as .pdf files. Version 1.23 of the Apache Tika Python package was used to convert the PDFs to raw text. From there, the texts were cleaned and “preprocessed.” This involved converting the text to all lowercase, removing spurious symbols (for example: “~”, “<”, “>”, “’”), and fixing common OCR errors like reading “©ur” instead of “our” or reading “ä” instead of “a”. Because NLP methods rely on punctuation to split sentences, common abbreviations that often create mistakes in the algorithms were removed. For example, all occurrences of “rev.” were substituted with the full word “reverend.”

Additionally, the first page or portions of the first page of many pamphlets were excluded because the OCR had difficulties with stamps and handwritten notes that were almost exclusively contained on the front covers. After text cleaning, the first 100 words of each text were isolated. Figure 1 shows 100 words of cleaned text extracted from “As An American” (Figure 2). The first 100 words serves as a rule-of-thumb baseline to score more advanced algorithms against. The reason this is a baseline is because the beginning of a text not only is easy to isolate, but often contains main ideas.



as an american it is easy to be satisfied with one's condition in life. despite the war and the immediate postwar problems, some men see in their imagination a peaceful security in the seclusion of their homes or in the rooms of an apartment house. from a local point of view there are some who are justifying their present contentment. in a way it is not difficult for one to say, and it has been said many times, "i am an american," and to be thoroughly satisfied with all that statement implies. in over a century and a half of

Figure 1. First 100 Words of the Pamphlet “As An American”

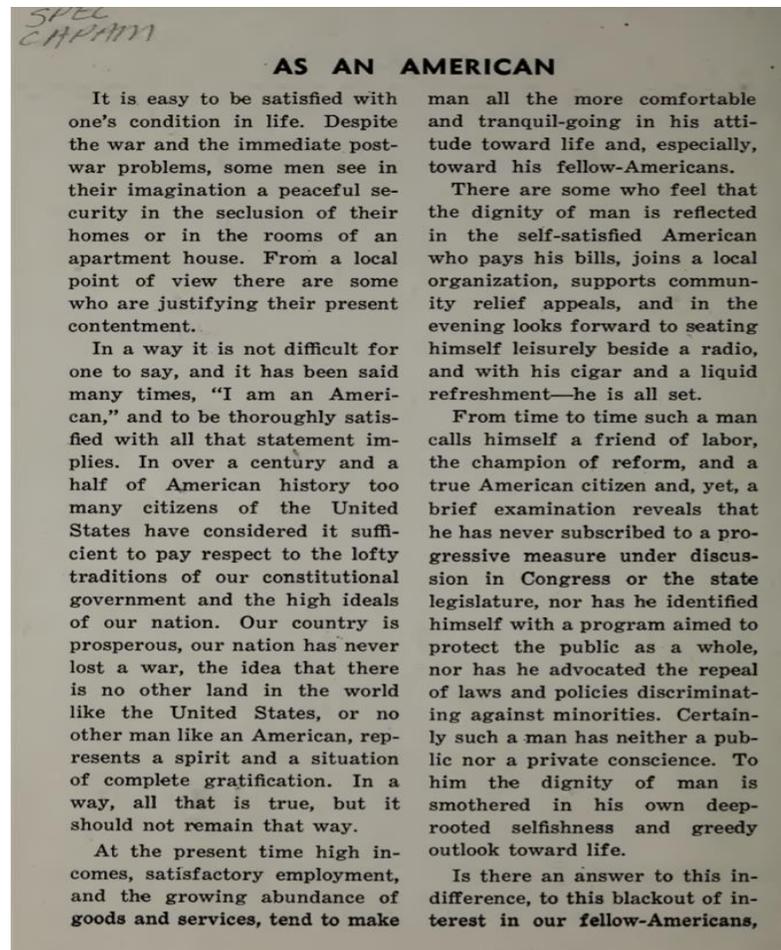


Figure 2. The First Page of the Pamphlet "As An American"

Some pamphlets contained sentences that were over 100 words. Often, summarization algorithms gave these long sentences too much weight. In such cases, sentences were split on commas if the tokenized text on each side of the comma was greater than eight tokens. Version 3.4 of the NLTK package (Bird, Loper & Klein, 2009) in Python was used to tag word labels (noun, verb, etc.) and to delete coordinating conjunctions following split commas. This preprocessing step improved output for the Textrank based algorithms, but this long sentence comma splitting step did not improve the output of the more sophisticated BERT algorithm.

### Summarization Algorithms

Four algorithms were used for the same summarization process. The result of each algorithm was evaluated and compared later on.

Extraction was performed on two to four-word phrases with the Rapid Automatic Keyword Extraction (RAKE) algorithm via version 1.4.5 of von Feilitzsch's python-rake package. The RAKE algorithm pinpoints keywords based on frequency and the "degree" of a word or phrase (Rose et al. 2010.) The degree of a word is based on many co-appearances a word has near other unique words. For example, if the phrase "credit union" appears in a sentence about bank

regulations and a sentence about Catholic credit unions, it would have a higher degree than if it only appeared in two sentences about Catholic credit unions. The python-RAKE package gives the option of setting a minimum frequency for extracted keywords, i.e., removing phrases from consideration if they do not appear in the document sufficiently many times. For two-word phrases, a minimum frequency of 5 co-occurrences was selected. Three-word phrases needed to have at least 4 co-occurrences. Four-word phrases needed to have at least 3 co-occurrences. There were no minimum requirements on the degree of words.

Two TextRank based algorithms were explored and used to extract summaries. TextRank is an NLP algorithm for keyword identification and summarization. It is based on the same principles as Google's PageRank algorithm. It ranks sentences that have word co-appearances in other sentences higher. It also takes into account the proximity of a keyword to other keywords (Mihalcea & Tarau, 2004). The first of the two was an algorithm the author hand-implemented based on instruction by Prateek Joshi (Joshi, 2018). This implementation can be called the raw TextRank implementation.

The second TextRank style of algorithm was Řehůřek's Python package Gensim. Version 3.8.0 of Gensim was used, which gives the option of setting a target word length. For this corpus of pamphlets, 75 words yielded better results than 50, 100, or 150 words. The hand-implemented TextRank algorithm was set to extract four-sentence summaries. The process plays out such that the algorithm would select words, phrases, and sentences from the existing corpus and then rearrange those elements into a (hopefully coherent) four-sentence summary.

Lastly, a recent, more state-of-the-art extraction method was applied: the Bert Extractive Summarizer (Miller, 2019.) This algorithm's eponymous Python package implementation, version 0.3.0 utilizes Google's BERT subject encodings and performs k-means clustering on sentences. It identifies sentences that are closest to the k-means centroid for the entire document. Unlike previous summarization methods, the BERT Extractive Summarizer utilizes deep learning. The model integrates Google's BERT (Bidirectional Encoder Representations from Transformers), a pre-trained set of subject encodings. By using Google's pre-trained encodings, this model forgoes the necessity to train on huge swaths of text data. The model instead benefits from BERT's previous machine-learned understanding of English texts when it performs clustering. The BERT Extractive Summarizer package's "ratio" option was used to extract and reassemble summaries that were roughly three sentences in length.

## Evaluation

The computer-generated summaries were benchmarked against their respective human-created reference summaries. Recall, approximately 101 of the 5500 pamphlets had acceptable, human-written summaries. The computer-generated versions of these same 101 pamphlets were compared and scored against their human counterparts. Here, the standard ROUGE (Recall-Oriented Understudy for Gisting Evaluation) F1 scoring test was used to compare the similarity of the summaries. Specifically, Tardy's rouge Python package, version 1.0.0, was imported to implement the comparisons.

ROUGE F1 is a common NLP metric for evaluating overlap between generated and reference summaries (Lin & Och, 2004). ROUGE F1 scores range between 0 and 1, with higher scores representing a greater percentage match between texts. The algorithm with the highest

average ROUGE F1 scores (across all 101 tests) would be deemed as best. For different types of documents, different ROUGE scores are acceptable. Because no research has focused on summarizing similar pamphlets, this project will help inform our perceptions of ROUGE scores.

## RESULTS

### ROUGE F1 Scores

RAKE-generated keywords were not scored, but they were kept as a possible addition to our library’s search layer. For illustrative purposes, Figure 3 contains an example of algorithm outputs for “As An American” Pamphlet. This is the same Pamphlet shown in Figure 1 - Appendix B shows the full text of the “As An American” pamphlet.

For the 101 pamphlets, the average ROUGE F1 score for the first 100 words vs. the human-generated summary was .109.

The hand-implemented TextRank algorithm performed better at an average of .144. The Gensim package’s implementation of TextRank yielded an average of .151. The BERT Extractive Summarizer performed significantly better than any previous method, with an average ROUGE F1 score of .239.

RAKE Keywords	TextRank	Gensim	BERT Extractive Summarizer
national broadcasting company, american citizen, catholic men, catholic hour, national council, united states public funds, george f Donovan president, make man, fair play, parochial school, public school, many americans, common welfare, nonpublic school, economic dignity, parents, freedom	as an american the catholic hour as an american by dr. george f. donovan president of webster college aug 28 946 the tenth in a series of addresses by prominent catholic laymen entitled the road ahead, delivered in the catholic hour . broadcast by the national broadcasting company in cooperation with the national council of catholic men, on august 11 , 1946 by dr. george f. donovan president of	the endorsement of a program involving a living wage for a man and his family, reasonable prices that will give a fair profit to the producer, and quality and service to the consumer, the extension of both public and private security to all americans, such as domestic workers, employees of nonprofit organizations, and rural workers, the recognition of collective bargaining, the abolition	as an american it is easy to be satisfied with one’s condition in life. the boss still secures silent and respectful obedience, and even the public gives adulation and praise to demagogues and to public leaders who have betrayed the confidence placed in them by the citizens. mere man, in this great maelstrom of political instability, is certainly not clothed in dignity.

Figure 1. Algorithm Outputs for "As An American"

## Subjective Analysis

Because there were two groups of reference summaries (summaries that were human written expressly to serve as reference summaries vs. the summaries imported from existing metadata for the same purpose), scores were differentiated within those groups. The 50 Pamphlets that had reference summaries generated for this project had better ROUGE scores (.179) across all methods when compared against the 51 summaries imported from our existing metadata (.143). Figure 4 shows how each algorithm performed against the two groups of reference summaries.

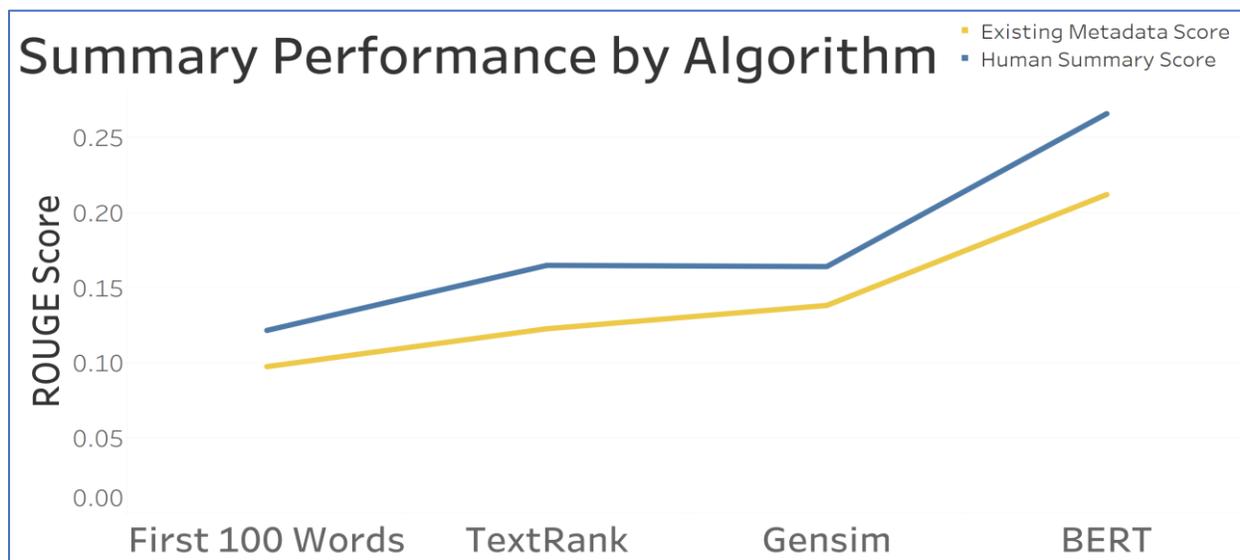


Figure 2. ROUGE Performance of Algorithms

Present in the summaries were also problems independent of whether they summarized a document well. Sentence fragments and unintelligible sentences, especially ones in which OCR erroneously inserted line breaks, frequently hampered all algorithms. For example, 14% of BERT summaries had large OCR errors present in them, and 6% had phrases in Latin. To understand these types of errors, examples from the BERT Extractive Summarizer are presented because BERT's results were deemed best by ROUGE scores and subjective analysis.

Early in the texts, the central ideas are more likely to appear, and the summarization algorithms weight location in text, which increases the likelihood OCR errors show up in summaries. OCR errors tend to happen toward the beginning of pages in headers or on the first pages in tables of contents and purchase notes. A heinous example of a front-loaded OCR error occurred in the BERT algorithm, which produced the following summary: "j& ociefg for fjfc jhopagftfiott of f § an international association for the assistance by prayers and alms of the missionary priests, brothers and nuns engaged in preaching the gospel in heathen and non-catholic countries. 5,510.60." After this OCR error crept into the automated summary, the rest of summary correctly gave the main idea of the pamphlet.

Full sentences in Latin are also distracting to users. The following summary generated by the BERT algorithm for a text on Catholic hymns is half in Latin: "words of catholic hymns.

congregation of holy cross the ave maria press notre dame, indiana 1924. iesu dulcis memoria, dans vera cordis gaudia sed super mel et omnia, eius dulcis praesentia." This auto-generated summary may give a patron some vague idea about the contents of its pamphlet (it includes Latin hymns), but unfortunately, any human-generated short summary would not include Latin. Thus, this example has a lower ROUGE score (.199) than the average for the BERT algorithm (.239).

## DISCUSSION

Scores of Pamphlets as a whole increased drastically from the baseline method of the first 100 words to the BERT-Extractive-Summarizer. Unfortunately, Latin text and OCR errors have persisted despite choosing more sophisticated algorithms and better text preprocessing.

For collections in which their text does not need OCR and all appears in English, the BERT Extractive Summarizer likely could provide summaries with few enough errors that catalogers would be comfortable adding the computed-generated summaries with minimal human oversight. In our corpus of historic Catholic Pamphlets, our institution will likely still need human review.

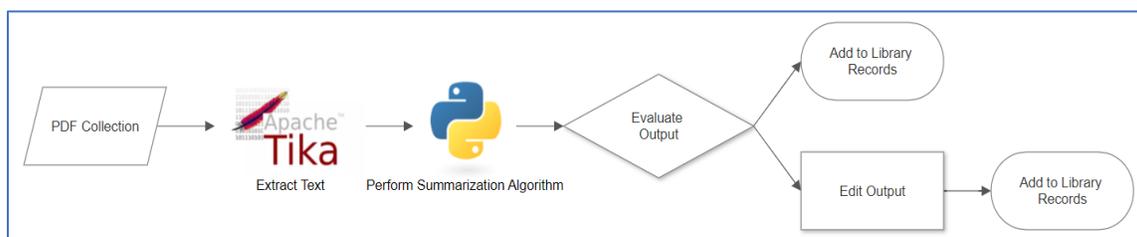


Figure 3: Example Workflow: Front-Facing Implementation

Our institution hopes to use these machine-generated summaries in two different capacities. First, the summaries may just be added to our catalogue “as is”, along with a note alerting the patron that the summary was machine generated. Second, staff may read through each machine-generated summary. The staff reviewer would then approve, slightly edit, or reject the summaries. This pipeline is illustrated in Figure 5. The first approach is more likely for our special collections website [curate.nd.edu](http://curate.nd.edu); the second is more suited to our MARC catalog.

### Future Improvements

Improving OCR issues is the biggest concern my library colleagues brought forward. In this project, open source software (Tika) was used to turn PDFs to text for reproducibility purposes. Repeating the whole pipeline using a paid OCR resource like Abbyy FineReader, which our library already pays to access, could yield more accurate raw text. If further OCR correction seems necessary, there are various NLP and statistical methods that have been researched for correcting OCR errors computationally (Amrhein & Clematide, 2018; Kumar, 2016). Excluding Latin text and sentence fragments is another non-trivial step. Fortunately, the Python NLTK package offers language detection (Bird, Loper & Klein, 2009), which can be used to exclude long segments of Latin Text. It also offers part-of-speech tagging, which can be used to exclude sentence fragments.

## Author Note

This project was supported by Jessica Morales and Laura Sill of Notre Dame Hesburgh Libraries' Resource Acquisitions and Discovery Program. PDF collections were supplied by Mikala Narlock with resources from the CurateND Support Team. This project grew out of a class project for Prof. David Chiang's Natural Language Processing class and was presented as a poster in March at Code4Lib Conference 2020 in Pittsburgh, Pennsylvania.

Please address questions and comments about this article to Jeremiah Flannery, 402, 284 Hesburgh Library, Notre Dame, IN 46556 or via email at [jflanne5@nd.edu](mailto:jflanne5@nd.edu).

## References

- Amrhein, C., & Clematide, S. (2018). Supervised ocr error detection and correction using statistical and neural machine translation methods. *Journal for Language Technology and Computational Linguistics (JLCL)*, 33(1), 49-76.
- Bird, S., Loper, E., & Klein, E. (2009). Natural language toolkit. *J URL Http://Www.Nltk.Org*,
- Borko, H., & Bernier, C. L. (1975). Abstracting concepts and methods.
- Chiang, D. (2019, November,). Lecture notes.
- CurateND Support Team, & McManus, J. (2019). The catholic pamphlets. Retrieved from <https://curate.nd.edu/show/1v53jw84s7d>
- Dong, Y., Shen, Y., Crawford, E., van Hoof, H., & Cheung, J. C. K. (2018). Banditsum: Extractive summarization as a contextual bandit. *arXiv Preprint arXiv:1809.09672*.
- Gupta, V., & Lehal, G. S. (2010). A survey of text summarization extractive techniques. *Journal of Emerging Technologies in Web Intelligence*, 2(3), 258-268.
- Joshi, P. (2018). An introduction to text summarization using the TextRank algorithm (with python implementation). Retrieved from <https://www.analyticsvidhya.com/blog/2018/11/introduction-text-summarization-textrank-python/>
- Kasten, S., & Flannery, J. (2020). *Bibliographic analysis of english language monographs*. Unpublished manuscript.
- Kumar, A. (2016). A survey on various OCR errors. *International Journal of Computer Applications*, 143(4), 8-10.
- Lin, C., & Och, F. J. (2004). Looking for a few good metrics: ROUGE and its evaluation. Paper presented at the *Ntcir Workshop*.
- Luhn, H. P. (1958). The automatic creation of literature abstracts. *IBM Journal of Research and Development*, 2(2), 159-165.
- Luk, A. T. (1996). Evaluating bibliographic displays from the user's point of view: A focus group study. *Master of Information Studies Research Project Report. Toronto: Faculty of Information Studies, University of Toronto*.
- Lundgren, J., & Simpson, B. (1997). Cataloging needs survey for faculty at the university of florida. *Cataloging & Classification Quarterly*, 23(3-4), 47-63.
- Lundgren, J., & Simpson, B. (1999). Looking through users' eyes: What do graduate students need to know about internet resources via the library catalog? *Journal of Internet Cataloging*, 1(4), 31-44.
- Mihalcea, R., & Tarau, P. (2004). Textrank: Bringing order into text. Paper presented at the *Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing*, 404-411.
- Miller, D. (2019). Leveraging BERT for extractive text summarization on lectures. *arXiv Preprint arXiv:1906.04165*.
- Parikh, A. P., Täckström, O., Das, D., & Uszkoreit, J. (2016). A decomposable attention model for natural language inference. *arXiv Preprint arXiv:1606.01933*.

- Rose, S., Engel, D., Cramer, N., & Cowley, W. (2010). Automatic keyword extraction from individual documents. *Text Mining: Applications and Theory, 1*, 1-20.
- Rush, A. M., Chopra, S., & Weston, J. (2015). A neural attention model for abstractive sentence summarization. *arXiv Preprint arXiv:1509.00685*.
- 

### **About the author**

Jeremiah Flannery is collections data analyst at University of Notre Dame's Hesburgh Libraries. He primarily works in statistical reporting and metadata quality. His research interests include machine learning, natural language processing, and the efficacy of metadata in relation to search layers and user experience.

## Appendix A: Benchmark Pamphlets

thecommitteeofco00nevi.pdf  
jointstatementoncath.pdf  
000757895.pdf  
nationalshrineof00nati.pdf  
apostolicasollicpaul.pdf  
tyrannyinmexicoscath.pdf  
christianityinte00nati.pdf  
eucharistsacramekoen\_0.pdf  
000758903.pdf  
letsworkwithgang00shaw.pdf  
thinkasyousayros00woma.pdf  
christchristians00doug.pdf  
prayersforfamily00schm.pdf  
asamerican00dono.pdf  
popepaulspeakstopaul.pdf  
asworldcitizen00dono.pdf  
statementofuscat00unit.pdf  
jesusireflection00null.pdf  
001008801.pdf  
statementonfeedi00unit.pdf  
003361303.pdf  
001008742.pdf  
instructiononfaccath.pdf  
decreeiissuedmotucath.pdf  
000757618.pdf  
interimguidelinecath.pdf  
apostolicexhortacath\_0.pdf  
000871207.pdf  
layapostolateitspius.pdf  
000735616.pdf  
storiesofgreatsa00trea.pdf  
holybibleheritag00cath.pdf  
000745365.pdf  
lawofcharity00mcme.pdf  
thisiscanabymr00farr.pdf  
authorizedinterp00sloa.pdf  
000832827.pdf  
wecanwemust00unse.pdf  
storyofnewcrusad00unse.pdf  
000746220.pdf  
000746204.pdf  
thirdinstructioncath.pdf  
000777747.pdf  
holinessspiritua00cath.pdf  
000746032.pdf  
000741054.pdf  
godlovessick00meeh.pdf  
wordstocatholich00cong.pdf

000626920.pdf  
001013628.pdf  
faithfulforlifem00cath\_0.pdf  
sacramentofmatri00harr.pdf  
001008065.pdf  
theycallitfreeen00smit\_0.pdf  
visittofrancisca00unse.pdf  
goodfridaydevoti00chur.pdf  
50000careerguida00unse.pdf  
000768904.pdf  
000833555.pdf  
universalnormofm00shee.pdf  
nonsectarianbus00mcma.pdf  
000769849.pdf  
suppressionofmon00unse.pdf  
whencecomewars00shee.pdf  
000998236.pdf  
000883213.pdf  
functionofartaddpius.pdf  
onafricapopepaulcath.pdf  
archbishoppurcel00hawk.pdf  
moralcurve00unse.pdf  
giveushearingstr00schr.pdf  
001006422.pdf  
000472199.pdf  
letterofjeancardvill.pdf  
001013833.pdf  
farmingasvocatiojohn.pdf  
popepaulonracismcath.pdf  
000762717.pdf  
ministeringtoadu00hugh.pdf  
howtoteachourlitschm\_0.pdf  
000997852.pdf  
002226254.pdf  
1942christmasmespius.pdf  
lamp1211soci.pdf  
lettertowomenreljohn.pdf  
seventythreecath00cath.pdf  
holybibleheritag00unse.pdf  
000757566.pdf  
howtoserve-lowmas00obri.pdf  
000746232.pdf  
mccarthyshistory00mcca.pdf  
reportoftrialofm00null.pdf  
administrationofcath\_0.pdf  
policystatementocath.pdf  
communitycrimest00unit.pdf  
statementofunite00unit.pdf  
000746248.pdf  
000828271.pdf  
youngsamornative00nati.pdf

000787820.pdf

birthcontrolunna00latz.pdf

## Appendix B: Raw Text of Pamphlet Titled “As An American”

"as an american it is easy to be satisfied with one's condition in life. despite the war and the immediate postwar problems, some men see in their imagination a peaceful security in the seclusion of their homes or in the rooms of an apartment house. from a local point of view there are some who are justifying their present contentment. in a way it is not difficult for one to say, and it has been said many times, "i am an american," and to be thoroughly satisfied with all that statement implies. in over a century and a half of american history too many citizens of the united states have considered it sufficient to pay respect to the lofty traditions of our constitutional government and the high ideals of our nation. our country is prosperous, our nation has never lost a war, the idea that there is no other land in the world like the united states, or no other man like an american, represents a spirit and a situation of complete gratification. in a way, all that is true, but it should not remain that way. at the present time high incomes, satisfactory employment, and the growing abundance of goods and services, tend to make man all the more comfortable and tranquil-going in his attitude toward life and, especially, toward his fellow-americans. there are some who feel that the dignity of man is reflected in the self-satisfied american who pays his bills, joins a local organization, supports community relief appeals, and in the evening looks forward to seating himself leisurely beside a radio, and with his cigar and a liquid refreshment—he is all set. from time to time such a man calls himself a friend of labor, the champion of reform, and a true american citizen and, yet, a brief examination reveals that he has never subscribed to a progressive measure under discussion in congress or the state legislature, nor has he identified himself with a program aimed to protect the public as a whole, nor has he advocated the repeal of laws and policies discriminating against minorities. certainly such a man has neither a public nor a private conscience. to him the dignity of man is smothered in his own deeprooted selfishness and greedy outlook toward life. is there an answer to this indifference, to this blackout of interest in our fellow-americans,

to this blindness toward the injustice practiced in our human relations. is there not a way to revive the dignity of man in the american scene. there is a reply which manifests itself in three areas political life, economic relations, and education. it does not take an overly astute ... student of american affairs to recognize the fraud and dishonesty practiced by some of our public officials both in high and low places. we all hear of the pressure exerted on members of our legislatures, of the offers and acceptances of bribes, of the lack of fairness in the administration of the law. the political machine is still known to be important. the boss still secures silent and respectful obedience, and even the public gives adulation and praise to demagogues and to public leaders who have betrayed the confidence placed in them by the citizens. mere man, in this great maelstrom of political instability, is certainly not clothed in dignity. he has lost practically all his own influence for the good of the common welfare. the answer to this problem is to be found in an enlightened and determined citizenry who have the courage to pursue the american way of fair play and justice to all citizens, irrespective of their race, creed, color, or status in life. the steps for reestablishing this impartial outlook are 1 an attitude of mind leading to the belief and the understanding that it is important to every american citizen to participate effectively and to the full in the public life of our country. 2 the establishment of a more complete and long-range program of education for our youth on the principles and practices of american government beginning with the town and going all the way up to the federal government itself, such a course should run from the elementary grade school through the college and university levels. 3 participation in movements designed to promote equality of opportunity for assuming responsibility in government among the people, particularly in the institution of such practices as the initiative, referendum, the recall, the extension of the merit system, greater scrutiny of public funds, and a more adequate expression of opinion on public

questions by the rank and file in government service. 4 the support of programs destined to spread democratic benefits among the citizens, especially in the removal of unfair voting restrictions, raising of the salary scale and promotional standards in public positions, and the expansion of educational and medical facilities throughout the nation. a lack of understanding of our government is manifested in the failure of many americans to recognize the inferior position of man in our economic life. the economic dignity of man has suffered much. many americans have not completely seen the necessity of fighting those evils which have threatened and lowered the economic status of many of our fellow-citizens. low wages, high prices, plus arbitrary control of prices and goods, strong monopolies, insecurity, especially in old age and for persons disabled by accident or illness, the great expansion of the apartments and hotels as the so-called substitute for the american home, the wide discrimination, especially in securing and holding a job, and conducting a business, are among some of the modern practices of american economic life that tend to reduce and break down the dignity of man. the answer to these economic problems is not hard to find. the endorsement of a program involving a living wage for a man and his family, reasonable prices that will give a fair profit to the producer, and quality and service to the consumer, the extension of both public and private security to all americans, such as domestic workers, employees of non-profit organizations, and rural workers, the recognition of collective bargaining, the abolition of racial, religious, and economic discrimination, greater freedom in the establishment of business, especially small business enterprises, greater opportunity in purchasing one's own home, are the answer. these solutions may be accomplished through a strong cooperative, mutual organization of all the groups concerned and if their efforts do not succeed entirely, or fail, then government has the obligation to step in for the purpose of providing for the common welfare of the distressed and handicapped citizens. as one who has had the experience of living in an indus

trial community and working in a factory, i recall very well one manufacturing concern owner who dedicated his life to the welfare of his employees and their families, not in a paternalistic role, but as a leader who recognized his essential obligations to his fellowman. his workers shared in the management, assisted in formulating policies, received a portion of the profits, took part in sales discussions and assumed many other obligations within the company. the result was very evident. high wages, year-round employment, satisfactory working conditions, friendly relations between the workers and the employers and never a strike in the entire history of the industry. an exception, yes, but it could easily be the rule. in addition to the political and economic dignity of man as factors in american success and happiness, there are also some educational implications which are pertinent. the united states has seen the growth of two systems of education—public and non-public. many of the non-public schools are controlled by the churches. both systems of education, public and nonpublic, train children for participation in american democratic life. under the constitution, american parents are assured of freedom of education. they may send their children to either a public school or a non-public school. in the celebrated oregon school case the supreme court of the united states ruled that no state law could abridge this fundamental american freedom by compelling parents to send their children only to the public schools. likewise, under our constitution, parents are guaranteed freedom of religion. they may worship god according to the dictates of their consciences. furthermore, they may choose for the education of their children a school where religion is taught along with the other subjects. any state or local law which would prohibit parents from enrolling their children in church controlled schools would be declared unconstitutional by our courts. although freedom of education and freedom of religion are guaranteed rights of every american citizen, in many cases american parents cannot exercise these rights. i know of a small community where the ma

jority of parents would prefer to send their children to a parochial school. this preference is based upon sound religious convictions. nevertheless, these parents are sending their children to a public school—why. the state will not provide any public funds for a parochial school. the parents and the other members of the religious group have low incomes. they simply cannot afford the cost of erecting and maintaining a parochial school in addition to paying high taxes for the public school. technically, therefore, these parents enjoy freedom of education and freedom of religion, but it seems to me these freedoms will have

little practical meaning in their lives unless the state adopts a fairer method of distributing public funds for educational purposes. the answer, simple and direct, according to traditional american practices, should be to give all parents the same opportunities to participate equally and fully in all educational plans and curricula so that the parent who desires to send his children to a non-public school where a belief in god is taught may do so without impairing his own financial status or depriving himself or his children of public funds which should be available to all americans. this program of equal distribution of public funds should begin on the local level so far as grade and high schools are concerned and on the state level for higher education. if necessary in larger areas, i believe the national government has a responsibility to step in and provide funds for students who wish to secure an education. this is fair play. it is the recognition of dignity in the parent who exercises the right to send his children to the school which his conscience selects. all through these areas, political, economic, and educational, where americans may and, should participate, there is ever present the idea of a permanent place of instruction from which the interest, the leadership, and the action should emerge. that place is the family. a stable, happy family where there are just and understanding relations between the husband and wife, between the father and the son, the mother and the daughter, where there is no room for unhappiness and where there will not be any divorce. such principles rooted in a true family life may be easily

translated into the political, economic, and educational fields where opportunities will be provided for the recognition of the obligations of the american citizen to his government, to his occupation, and to the education of his children. god in his greatness and kindness will then give, as he does to the family, that love of fellowman, a love which respects the conscience of man, which stresses the loyalty of man to the state through citizenship and which permits economic cooperation among groups and which will make man a better american citizen. through these means the dignity of man will be appreciated and the citizen of these united states will be more proud than ever before to say, "i am an american." 4

j 3 2223 the catholic hour 1930—seventeenth year—1946 the nationwide catholic hour was inaugurated on march 2, 1930, by the national council of catholic men in cooperation with the national broadcasting company and its associated stations. radio facilities are provided by nbc and the stations associated with it the program is arranged and produced by nccm. the catholic hour was begun on a network of 22 stations, and now carries its message of catholic truth on each sunday of the year through a number of stations varying from 90 to 110, situated in 40 states, the district of columbia, and hawaii. consisting usually ' of an address, mainly expository, by one or another of america's leading catholic preachers—though sometimes of talks by laymen, sometimes of dramatizations —and of sacred music provided by a volunteer choir, the catholic hour has distinguished itself as one of the most popular and extensive religious broadcasts in the world. an average of 100,000 audience letters a year, about twenty per cent of which come from listeners of other faiths, gives some indication of its popularity and influence. our sunday visitor press huntington, indiana

.."