

**Puzzle Poetry**  
(Technical Paper)

**Ethics of Data Collection**  
(STS Paper)

A Thesis Prospectus Submitted to the  
Faculty of the School of Engineering and Applied Science  
University of Virginia • Charlottesville, Virginia  
In Partial Fulfillment of the Requirements of the Degree  
Bachelor of Science, School of Engineering

Daniel Patel  
Fall 2019

Technical Project Team Member  
Kaan Katircioglu

On my honor as a University Student, I have neither given nor received  
unauthorized aid on this assignment as defined by the Honor Guidelines  
for Thesis-Related Assignments

Signature **Daniel Patel** Date 4/27/2020  
Daniel Patel

Approved \_\_\_\_\_ Date \_\_\_\_\_  
Nathan Brunelle, Department of Computer Science

Approved \_\_\_\_\_ Date \_\_\_\_\_  
Richard Jacques, Department of Engineering and Society

## **Introduction:**

Throughout history, one of the most influential factors in the rise and fall of civilizations was access to information. For example, The Caesar Cipher allowed Julius Caesar to be victorious in ways he otherwise would not have been because his enemies never had access to his plans. Similarly, in WWII, the allies were able to shorten the war significantly once they were able to gain access to Nazi instructions by means of finding the necessary information to decode messages encoded by enigma machines. In light of this, it seems of relative importance to reflect on the nature of this valuable resource in the present context. In modernity, information plays as pivotal a role, yet is far easier to obtain than it has been at any point in history. Through the use of the internet, cellphones, and their respective networks, email, and cameras and microphones installed in many devices, it is evident that nearly everything people do is recorded and saved (“Privacy Policy – Privacy & Terms – Google,” n.d.). Because of this, there is necessarily nearly unfathomable amounts of information about everyone who partakes in the technological revolution of the past two decades. Because of the existence of this data, and the seeming relevance of it to social change, it becomes apparent that an exploration into the preferred allowance of the dissemination thereof should be undertaken. Specifically, the Research portion of my paper shall address the nature of the collection and usage of data by Google, a large player in the modern technological explosion, and the technical portion will be dealing with a problem created by an overabundance of data. In doing this, the aim of the paper is to illuminate the unique ethical issues brought up by the rapid increase in available information.

**Technical Topic:**

Brad Pasanek, a professor of English at the University of Virginia came up with the idea of turning sonnets into puzzles by partitioning the syllables of the poems into tiles. He was able to accomplish this easily with the lower portion of the sonnet but was unable to do so with the upper portion due to the enormous size of the search space. It is presumed that his algorithm's efficiency was greatly hindered because it stored too many states in memory, and, in doing so, caused the computer it ran on to have to swap portions of memory between the hard drive and RAM, a task that could likely slow a problem down by six orders of magnitude.

Because he was unable to devise an algorithm to tile poems such that the tiles did not split words, he enlisted the help of Computer Science students to attempt to crack the problem. The goal of the research is to develop an algorithm that can tile an arbitrary shape given a set of tiles such that the placement of the tiles does not cause constraints, given in the form of multiple syllable words, to be split into multiple tiles. From this, secondary objectives are to apply it to tile both the top eight and bottom six lines of a sonnet as well as to find reductions of the problem to and from known NP Complete problems, proving it to be complete. The research will be completed over the course of the Spring 2020 semester and both Kaan Katircioglu and Daniel Patel will be working together to attempt the creation of such an algorithm.

**STS Topic:**

In the present day, many people would be shocked to know the extent to which companies such as Google collect information about them. For instance, Google collects phone numbers, device information, every URL request made to Google by a user as well as the date and time thereof, the apps downloaded on users phones, videos watched, interaction with ads, YouTube comments, information about the voice of a user, purchase activity, who the user

communicates with, every site a user visits on chrome, location via GPS data, user IP addresses, information collected from user device sensors, and which cell towers, Wi-Fi networks, and Bluetooth devices connected to by user devices, billing information about users, as well as many other data points about their users (“Privacy Policy – Privacy & Terms – Google,” n.d.; “Google Maps Platform Terms of Service | Google Maps Platform | Google Cloud,” n.d.; “Data Transparency | Google Safety Center,” n.d.; Android – Android Enterprise Terms,” n.d. -a). Similarly, Google acquires the rights to process, redistribute, collect usage statistics, and provide to third parties this personal information (“Google Play Terms of Service,” n.d.; “Terms and Conditions,” n.d.; “Terms of Service – YouTube,” n.d.; “YouTube Data Processing Terms – YouTube,” n.d.)

Given this enormous, unprecedented amount of information a company has about its users, it seems appropriate to analyze the nature of the collection and usage of this information in an attempt to determine if the process should be considered moral, and from that, whether it should be allowed by society.

To approach the ethical nature of this problem, an ethical framework must first be adopted. In the paper, the framework employed will be a branch of deontological ethics derived from the Kingdom of Ends Formulation of Kant’s Categorical Imperative. Kant argues that both each formulation of the Categorical Imperative is equivalent and that each provides a consistent guide by which the morality of a given action can be assessed (“The Categorical Imperative,” n.d.). This specific formulation of The Categorical Imperative states that to be moral one must “act in accordance with the maxims of a member giving universal laws for a merely possible kingdom of ends” (Johnson & Cureton, 2019). By a kingdom of ends, Kant is referring to the notion of an ideal state consisting of the union of the idealized version of a person, the rational

being, through a shared set of laws (Bennet, n.d.). This formulation implies that morality is acting in accordance with the laws that would be set in place by a perfect legislature of rational entities governing a group of rational entities from which it is comprised. Based on this principle, for Google's actions to be considered moral, they must be such that a rational entity would allow such actions to be taken towards themselves.

In addition to this moral framework, the STS framework of technological momentum will be relied upon to argue both the necessity of addressing this question in a timely manner as well as the efficacy of doing so. When addressing this question, the stakeholders are both the users of Google's services as well as those who benefit from the data collected by Google such as advertisers. The primary artifact is the information collected, and other artifacts include the physical devices which generate and send data to Google. Given these actors, and this framework, it can be considered whether the actions of the company in question should be considered moral.

### **Research Question and Methods:**

The primary question to be addressed in this paper will be that of whether the methods employed by Google regarding the collection and usage of data should be considered moral. This question is of great importance in modernity due to the pervasive nature of the specific company in society. To accomplish this, the primary method to be employed will be a derivative of policy analysis. The policy analysis will be unconventional in that it will focus on corporate rather than governmental policy but in practice will have similar effect due to the relative power of the corporation in society. With this adjustment, the method serves as a helpful tool in addressing the

problem in that it provides a way to systematically lay out the background information in a way that the proposed moral framework can interact with.

The structure of the paper will be composed of three major parts. First, the policy will be collected and examined. After this, the moral framework will be employed to conjecture about an ideal, and then infer the necessary implications of this ideal. These implications will then be applied to the findings of the policy analysis in an attempt to see if the policy falls in line with the musings, resulting in the conclusion to whether or not the actions in question fall in line with the framework being employed.

### **Conclusion:**

Based on the relative importance of information in societies, it is clear that thought must be invested by these societies into their collective approach to the information. In the technical portion of the paper, a means by which information unnecessary to finding solutions must be determined and discarded will have to be found and employed. The desired result of this research will be a program which can take in an arbitrary poem and partition it into a set of tiles. In the STS section of this paper, a subset of internet data collection is observed in an attempt to judge whether it should be considered moral. This section will conclude whether the actions of Google with regards to data should be considered immoral based off a subset of deontological ethics. This conclusion will offer a perspective by which society could look at a specific company, and in doing so, could offer reason for the society to partake in the crafting of legislation to curtail activities deemed inappropriate by the company, under the assumption that society may deem such actions necessary. In any case, it is clear that, to this point, the pace at which information can be collected has been greater than the pace at which society has decided how it should place

bounds on this collection, and therefore, it seems necessary for society to further partake in this latter action.

## References:

Android – Android Enterprise Terms. (n.d.-a). Retrieved October 1, 2019, from Android website:

<https://www.android.com/enterprise/terms/>

Android – Android Enterprise Terms. (n.d.-b). Retrieved October 1, 2019, from Android

website: <https://www.android.com/enterprise/data-protection/terms/>

Bennett, J. (n.d.). Groundwork for the Metaphysic of Morals. *Immanuel Kant*, 53.

Data transparency | Google Safety Center. (n.d.). Retrieved October 1, 2019, from

<https://safety.google/privacy/data/>

Google Maps Platform Terms of Service | Google Maps Platform | Google Cloud. (n.d.).

Retrieved October 1, 2019, from <https://cloud.google.com/maps-platform/terms/>

Google Play Terms of Service. (n.d.). Retrieved October 1, 2019, from

<https://play.google.com/about/play-terms/index.html>

Johnson, R., & Cureton, A. (2019). Kant's Moral Philosophy. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2019). Retrieved from

<https://plato.stanford.edu/archives/spr2019/entries/kant-moral/>

Privacy Policy – Privacy & Terms – Google. (n.d.). Retrieved October 1, 2019, from

<https://policies.google.com/privacy?hl=en#infocollect>

Terms and conditions. (n.d.). Retrieved October 1, 2019, from Android Developers website:

<https://developer.android.com/studio/terms>

Terms of Service—YouTube. (n.d.). Retrieved October 1, 2019, from

<https://www.youtube.com/static?template=terms>



The Categorical Imperative. (n.d.). Retrieved October 1, 2019, from

<https://www3.nd.edu/~rbarger/categorical-imperative.html>

YouTube Data Processing Terms—YouTube. (n.d.). Retrieved October 1, 2019, from

[https://www.youtube.com/t/terms\\_dataprocessing](https://www.youtube.com/t/terms_dataprocessing)