CUNY-IBM Watson Case Competition 2018

Team Number: 53

Team Name: CSI Devs

Business Case Title: Synthia: The SMS Census Assistant

Link to the 1-minute: https://www.youtube.com/watch?v=ZpCyj5fSHcI

Sector: City Services

One-line about your project: New Yorkers will have the option to text an AI named Synthia answers to the 2020 Census along with any questions they have.

Team Members: Egor Semeniak, Anthony Astarita, Yuri Yurchenko, Vincent Vitiello

Summary: How Big is the Big Apple?

Every New Yorker needs to be counted. Period. Every single New Yorker deserves the representation, the funds, and the recognition they deserve. The United States Census gives the city its chance to be rightfully accounted for. Only one problem: too many New Yorkers don't get counted. According to an article from the New York Times, the mail participation rate for New York City was an astonishing 59 percent for the last decennial Census. The same article goes on to say this measly 59 becomes overshadowed by the national average mail-back rate of an overshadowing 71 percent.[1]

Our solution: a helpful AI known as Synthia. Citizens will receive a phone number on their Census form that they can text–at their convenience–to speak with Synthia. Synthia will text them the same questions found on the Census and send the answers directly to the government. Synthia can speak in any language, answer any questions they may have, all while keeping their data safe and secure.

1. The Market: Who does Synthia Help?

a) The average busy New Yorker

Synthia is leaps and bounds more convenient than mailing paper. New Yorkers are some of the busiest people on this planet. Some New Yorkers might not want to go through the hassle of mailing back the Federal Census. Instead, it would be far more convenient to text Synthia while in line for Starbucks or waiting for the next R train. According to a study by the United States Census Bureau, SMS messaging had an 80% response rate among the participants in their survey. In addition, the researchers found that 92% perceived the survey as "easy to complete." [2] If New York City adopted an SMS messaging approach to the Census and received similar results to this experiment, then the city would enjoy a 21 percent leap in Census participation and the subsequent increase in federal funds that would follow.

The lack of Census participation might be in large part because people don't mail the Census back. According to one Gallop/CNN poll, "data suggest that practical reasons -- including not receiving the form to begin with, or simply forgetting to mail it back in -- may be the biggest culprit." [3] No matter where a New Yorker is, as long as they're carrying their mobile phone, they can talk with Synthia. New Yorkers would not need to worry about remembering to mail back the Census.

b) Synthia targets homeless New Yorkers

Accounting for the homeless of New York City has consistently been a challenge of the Census Bureau. The current method of counting the homeless population involves having field

---

[1] Roberts, Sam. "New York City's Response Rate in Census Improves." *The New York Times*, The New York Times, 21 Apr. 2010, www.nytimes.com/2010/04/22/nyregion/22census.html.

[2] . Hicks, Wendy, Brett McBride, Krista Freedman, and Max Lin. "Exploring SMS as a Data Collection Method." Census.gov. Accessed April 20, 2018. https://www.census.gov/fedcasic/fc2010/ppt/06_hicks.pdf.

[3] Newport, Frank. "Very Few Americans Have Personal Objections to Filling Out Census Form."*Gallup.com*, 13 Apr. 2000, news.gallup.com/poll/3004/very-few-americans-personal-objections-filling-census-form.aspx.

workers go from shelter to shelter making estimates of the population, along with searching streets for the homeless. How Synthia assists the homeless of NYC is by putting the headcount in their own hands. A study performed by the Keck School of Medicine of University of Southern California found the vast majority of homeless citizens can "73% of total participants said that their cell phone can both send and receive text messages." This would allow roughly 70% of the homeless population to becomes accounted for without any external help.[4] This cuts the middle man field worker from guessing how many homeless people there are in NYC simply from select shelters.

2. The Solution: How Does Synthia Work?

a) Safely sends answers to government

The process begins when the user texts Synthia from the number provided on their Census form. The participant will then send a message to a pre-set Twilio number. The message is sent via HTTPS protocol, securely and encrypted. That message from Twilio gets sent to the Census Bureau's Node.js server. The Node.js server is running a custom script to work with up to 10,000,000 users at the same time. When their message is decrypted on the server, it gets sent as an API call to Watson Assistant. The process is that custom set on a server, based on the response from the Watson. All data stored on our server is recorded via our custom created encryption algorithm with two factor encryption that presents maximum security to the user. Data gets sent directly to the government, without any middleman holding the data.

b) Answers any questions citizens may have

Synthia has the ability to answer any questions the public might have about the Census. Instead of using Google or 311 all someone needs to do is ask to Synthia. Many people might be confused with a question on the Census, put the form down, and forget it ever exists. Now with a personal Census assistant, their question can be answered immediately. For example, the participant could ask Synthia something along the lines of "Will my answers be confidential?" to which Synthia could explain the laws, fines, and safeguards put in place to protect their privacy.

c) Accommodates to the Immigrant Population

IBM Watson's Natural Language Understanding allows Synthia to process thirteen different languages. Once a citizen texts the given number on their Census form, Synthia will only ask and answer questions in that language unless otherwise specified. As a city of immigrants, it is necessary Synthia accommodates all New Yorkers. In addition to assisting a diverse language spectrum, Synthia will also accommodate those not wishing to answer a particular question. For example, if there is a question regarding citizenship, Synthia will be sure to allow any participant to skip that question without penalty. Synthia's goal is to increase Census participation, and any question that might hinder that goal could be skipped.

---

[4] Rivera, Jordan. "Mobile Phone and Internet Access Among Homeless and Low Income Populations." Nmfonline.org. August 1, 2014. Accessed April 20, 2018. https://nmfonline.org/wp-content/uploads/2016/02/Rivera-Jordan-Paper.pdf.

3. Cost Benefit Analysis: What can Synthia Save NYC?

The Census becomes less expensive the more people respond. "It costs us just 42 cents in a postage paid envelope when households mail back their … forms [and] about $25 per person if we have to go out and knock on the doors," Robert M. Groves, the Census Bureau's then-director, said in a 2010 statement. The Census Bureau employed 635,000 workers in total for the 2010 Census. After the initial mailing of the 2010 Census, the Bureau sent thousands of workers door to door in NYC to try to find information about the participants living in households that did not mail the Census back. As stated in a NY Times article, "In 2010, the final mail-return rate for New York City was 63 percent." [6] Meaning, after the government spent $25 dollars per person, they only gathered information about 4 more percent of NYC. We predict Synthia will be able to replace at least 30% of every Census worker they deployed in 2010. This due to the estimated improvement using SMS as a Census alternative, along with the other aforementioned benefits to using Synthia. A better response rate will mean less people are needed to knock door to door to get information from NYC residents.

Action:	Description	Amount
Total cost of text by people to fill out Census:	Cost per text ($0.001) * Minimum amount of text messages sent (30) * Number of households	$93,000
Cost of Watson server per state:	Only counting for use in NYC	$400.00
Total cost of texts census:	Cost of paper + cost of text + Watson server	$93,400
Current cost to send field workers to houses	Number of households that did not mail back 2010 Census (961,000) * $25 per house	$24,025,000
Cost with only 70% of field workers	Current cost to send field workers to houses*0.7	$16,817,500
Total Savings:	Current cost to send field workers to houses subtracted from cost with only 70% of field workers	$7,207,500

References

5. U.S. Census Bureau 2010. *Fast Facts 2010* https://www.census.gov/history/www/through_the_decades/fast_facts/2010_fast_facts.html

6. Barron, James. "Preparing for the 2020 Census, One Address at a Time." *The New York Times*, The New York Times, 9 Mar. 2018, www.nytimes.com/2018/03/09/nyregion/census-2020-new-york.html.