# Assignment 1

| | |
|---|---|
| Created | @August 30, 2022 6:07 PM |
| Class | CSP 524- Big Data Technologies |
| Type | Assignment |
| Materials | |

Part 1 & 2:

1. What was the problem with the Google flu detection algorithm?

   GFT was designed to record the influence illness. However, the record detection between CDC and GFT was different because GFT was predicting the record count based on the internet search whereas CDC was tracking the records based on the patient's visit to the hospital. This raised an issue how far the algorithms being useful compared to the traditional method of tracking. Two main factors that created this issue are big data hubris and algorithms dynamics.

2. What is big data hubris?

   It is an assumption that big data are a replacement to the traditional method of collecting data. However, big data don’t assure the validity and reliability of the data. The main issue in the Big data hubris is, the methods/mediums it used to collect data are not a reliable source for scientific analysis.

3. What approach could have been used to improve the Google flu detection algorithm?

   Instead of using only the data collected from search results, if GFT data can be combined with the real time data acquired through traditional methods. And comparing the data, GFT could recalculate its results which could have provided more reliable output.

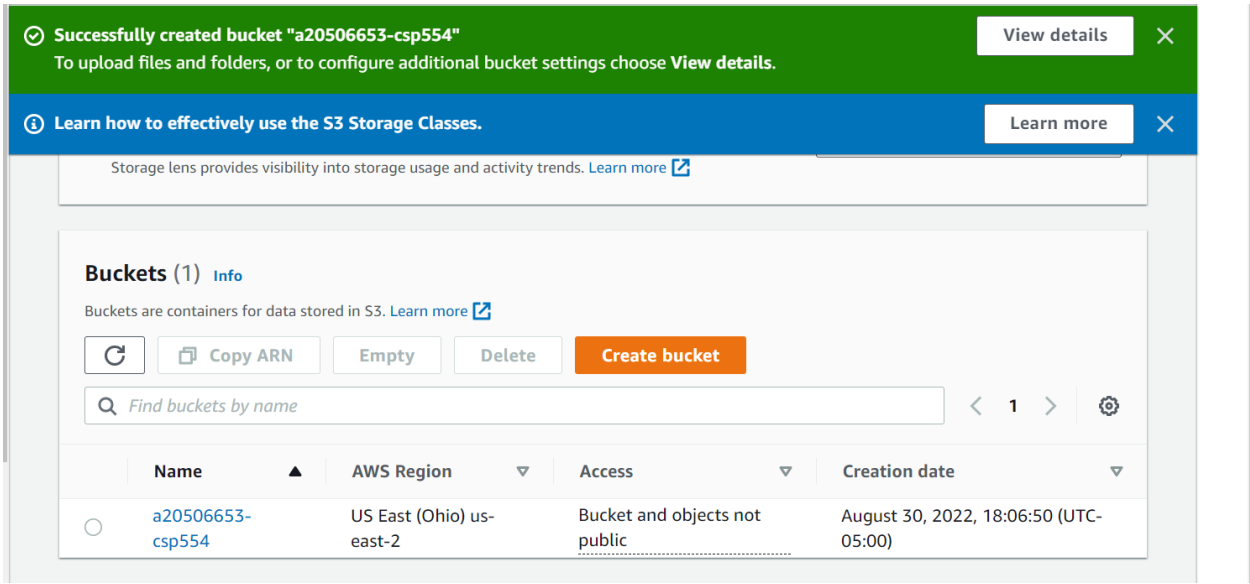
4. What is “algorithm dynamics?”

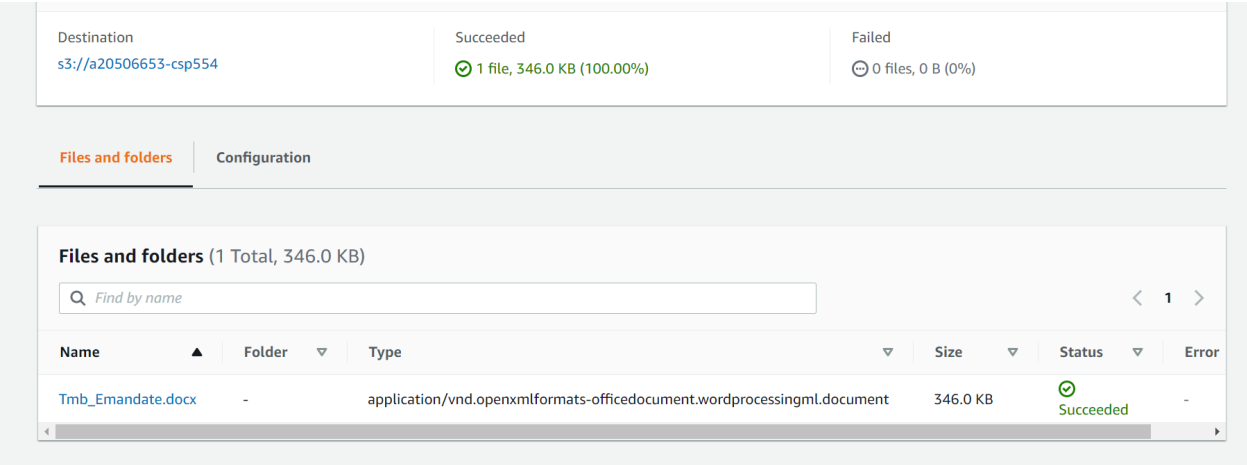
   It is the alterations that engineers make a commercial service as well as those that customers make while using it.  A number of modifications in GFT's tracking were probably impacted by the Google search algorithm, and user activity.

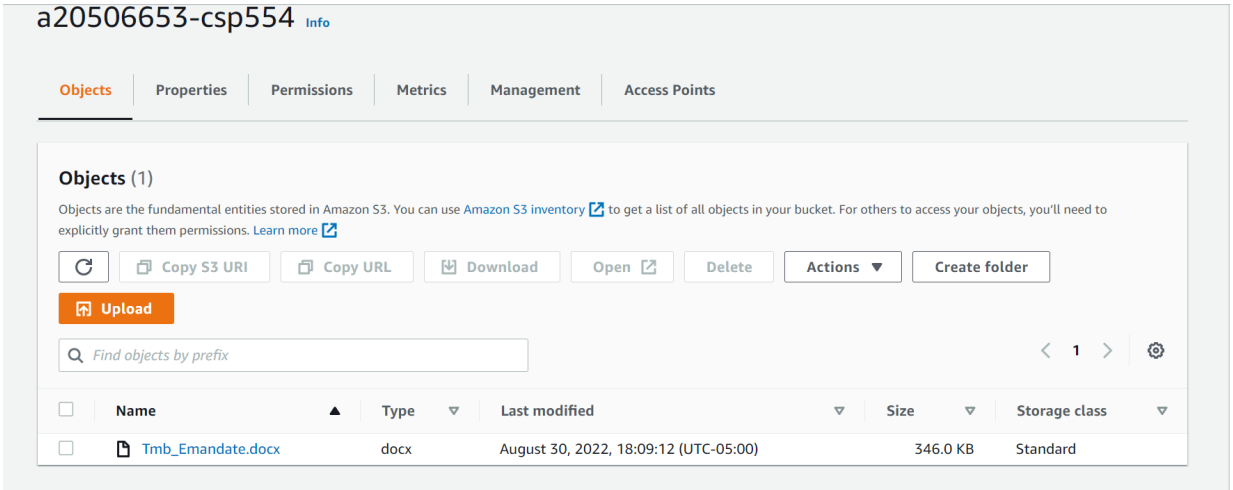
5. What aspect of algorithm dynamics impacted the Google flu detection algorithm?

   The company keeps updating the algorithm with increased search patterns which results in wide search results which could be unrelated and impacted the GFT algorithm. The search patterns are the decisions of many programmers across the organization and then the customers.

Part 3:




a20506653-csp554 Info
Objects
Properties
Permissions
Metrics
Management
Access Points
Objects (1)
Objects are the fundamental entities stored in Amazon S3. You can use Amazon S3 inventory to get a list of all objects in your bucket. For others to access your objects, you'll need to explicitly grant them permissions. Learn more
Copy S3 URI
Copy URL
Download
Open
Delete
Actions
Create folder
Upload
Find objects by prefix
1
Name
Type
Last modified
Size
Storage class
Tmb_Emandate.docx
docx
August 30, 2022, 18:09:12 (UTC-05:00)
346.0 KB
Standard