

Google-NYT project forecast: Cloud-y with bursts of visual impact

November 13 2018, by Nancy Owano



Credit: CC0 Public Domain

A newspaper morgue cannot afford to be a place where valuable photos

exuding barely told stories and teaching moments in history are dead and buried forever. A newspaper, after all, is a story-teller and *The New York Times* is out to make sure it leverages over 5 million items to keep the stories coming, with context, impact and teachable moments.

Nancy Weinstock, former photo editor, recalled for every picture they were able to publish many never saw the light of day. After all, it can be said that telling stories in pictures is the way society works now. More visually rich stories is apparently the key way to go.

Google likes that kind of pursuit. What a way to brand its Google Cloud expertise, from help wanted, to help at your service

The New York Times archive is in a project initiative with Google Cloud, which is helping the news mavens digitize millions of photos from its archive. Google Cloud posted the video about the undertaking on Nov. 9.

"..First New York City subway..and there you have it." Jeff Roth, archive caretaker, TNYT, is in the video, showing the world some of the many, many cabinets and drawers housing the photographs dating from the late 1800s. The Morgue is one reminder of the pivotal role the paper has played in newspaper journalism getting stories out through the years.

Cloud supporters may have reason to be enthused that the project takes over 100 years of photos in a basement and makes use of its treasures, which in turn allows people to see photos that they may have never been seen before, and makes them accessible to the newsroom.

"Digitize" barely reflects the technical feats needed to pull this off for the 5 million to 7 million items involved. They are not all same size entities, and that is just part of the situation; there are also those prints and contact sheets showing all the shots on photographers' rolls of film. Tyler Lee in *Ubergizmo* said "This also means that in the future trying to

access these photos will be easier as they will be searchable based on how they have been archived using Google's AI system."

On CNET, Stephen Shankland said, "The Times is using Google's technology to convert it into something more useful than its current analog state occupying banks of filing [cabinets](#)." Shankland crystallized what is going in here: "Neural networks are analyzing photos and captions dating back to the [1870s](#)."

What kind of technical process goes to work with Google Cloud now in the picture?

Google's Sam Greenfield walked readers through the process in a blog posting. This is described as part of the process: An image is ingested into Cloud Storage. The Times uses Cloud Pub/Sub to kick off the processing pipeline to accomplish tasks. Images are resized through services running on Google Kubernetes Engine (GKE) and the image's metadata is stored in a PostgreSQL database running on Cloud SQL.

The [images](#) are stored in Cloud Storage multi-region buckets for availability in multiple locations.

Tyler Lee observed a payoff in NYT's use of Google's tools. "Companies digitizing physical copies of photos and documents isn't new, but with the use of Google's AI, not only can this information be [scanned](#), but text accompanying the images will also help provide context about the photo, which in turn allows them to be catalogued and sorted automatically."

All in all, Greenfield described three contributions that Google Cloud brings to the NYT table: (1) they can securely store the images (2) give them a better interface for finding photos (3) they can find new insights from data locked on the backs of images. Google hopes that they can inspire more organizations—not just publishers—to look to the cloud,

and tools like Cloud Vision API, Cloud Storage, Cloud Pub/Sub, and Cloud SQL, to preserve and share rich histories.

More information: [cloud.google.com/blog/products...s-of-archived-photos](https://cloud.google.com/blog/products/ai-machine-learning/s-of-archived-photos)

© 2018 Science X Network

Citation: Google-NYT project forecast: Cloud-y with bursts of visual impact (2018, November 13) retrieved 19 April 2024 from <https://techxplore.com/news/2018-11-google-nyt-cloud-y-visual-impact.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.