

MISR.SAUDER.UBC.CA

MDS CAPSTONE

Extracting a Corporate Social
Network from SEC Filings

March 1, 2019



a place of mind
THE UNIVERSITY OF BRITISH COLUMBIA



AGENDA

- Who we are
- What we do
- What we want
- What data we have
- What we expect

WHO WE ARE

- MANAGEMENT INFORMATION SYSTEMS RESEARCH
 - Accounting and Information Systems Division
 - UBC Sauder School of Business
- Faculty Members
 - Hasan Cavusoglu
 - Gene Moo Lee

- Ph.D. Students:
 - Ozum Nazli Kafaee
 - Myunghwan Lee
 - Moksh Matta
 - Amir Bahreini
 - Ruijing Zhao
- Master Students
 - Victor Song
 - Rui Cao
- Undergrad Students
 - Christina Cheung
 - Raymond Situ



HASAN CAVUSOGLU

- Associate Professor
- B.Sc. Electrical and Electronics Engineering (Bogazici, 99)
- M.Sc. Telecommunications Consultancy (UT Dallas, 03)
- Ph.D. Management Science (UT Dallas, 03)
- Teaching:
 - Strategic Management of IT, Project Management, Database Management
- Research: Economics of Information Systems
 - Game Theory, Econometrics, Social/Behavioral Science methods

GENE MOO LEE



- Assistant Professor
- B.Sc. Computer Science & Mathematics (Korea U.)
- M.A. Computer Science (UT Austin, 06)
- Ph.D. Computer Science (UT Austin, 15)
- Industry Experiences: Samsung, Goldman Sachs, Intel, AT&T
- Teaching:
 - Business Analytics, Data Science for Business
- Research:
 - Methods: Machine Learning, Text Mining, Deep Learning
 - Topics: Industry Intelligence, Online Platforms, Social Media, Cybersecurity

WHAT WE DO

- Economics
- Econometrics
- Applying ML and AI into business research

WHAT WE WANT

- Constructing **Social Network of Public Corporations**
- Extract and visualize **inter-corporation relationships** disclosed by public companies in their **annual reports** to the U.S. Security and Exchange Commission (**SEC**)
- Delivered in the form of **Data Science Pipeline** and **Visualization**

WHAT WE HAVE

- Firm disclosure data from SEC EDGAR database
 - 11.8M filings, 736 types, 558K entities (firms, executives, and investors) in 24K cities (as of Feb 27, 2019)
 - 164K 10-K Annual Reports
- Firm performance data via WRDS (financial & accounting numbers)
 - Stock prices (CRSP)
 - Wharton Research Data Services (Compustat)

ALPHABET INC. 2018 FORM 10-K

<https://www.sec.gov/Archives/edgar/data/1652044/000165204419000004/goog10-kq42018.htm>

UNITED STATES
SECURITIES AND EXCHANGE COMMISSION
Washington, D.C. 20549

FORM 10-K

(Mark One)

ANNUAL REPORT PURSUANT TO SECTION 13 OR 15(d) OF THE SECURITIES EXCHANGE ACT OF 1934

For the fiscal year ended December 31, 2018

OR

TRANSITION REPORT PURSUANT TO SECTION 13 OR 15(d) OF THE SECURITIES EXCHANGE ACT OF 1934

For the transition period from _____ to _____

Commission file number: 001-37580

Alphabet Inc.

(Exact name of registrant as specified in its charter)

Delaware

(State or other jurisdiction of incorporation or organization)

1600 Amphitheatre Parkway

Mountain View, CA 94043

(Address of principal executive offices) (Zip Code)

(650) 253-0000

(Registrant's telephone number, including area code)

Securities registered pursuant to Section 12(b) of the Act:

Title of each class
Class A Common Stock, \$0.001 par value
Class C Capital Stock, \$0.001 par value

61-1767919

(I.R.S. Employer Identification No.)

Name of each exchange on which registered
Nasdaq Stock Market LLC
(Nasdaq Global Select Market)
Nasdaq Stock Market LLC
(Nasdaq Global Select Market)

Securities registered pursuant to Section 12(g) of the Act:

Title of each class
None

Indicate by check mark if the registrant is a well-known seasoned issuer, as defined in Rule 405 of the Securities Act. Yes No

Indicate by check mark if the registrant is not required to file reports pursuant to Section 13 or Section 15(d) of the Act. Yes No

Indicate by check mark whether the registrant (1) has filed all reports required to be filed by Section 13 or 15(d) of the Securities Exchange Act of 1934 during the preceding 12 months (or for such shorter period that the registrant was required to file such reports), and (2) has been subject to such filing requirements for the past 90 days. Yes No

Indicate by check mark whether the registrant has submitted electronically every Interactive Data File required to be submitted pursuant to Rule 405 of Regulation S-T (§232.405 of this chapter) during the preceding 12 months (or for such shorter period that the registrant was required to submit such files). Yes No

Indicate by check mark if disclosure of delinquent filers pursuant to Item 405 of Regulation S-K (§229.405 of this chapter) is not contained herein, and will not be contained, to the best of the registrant's knowledge, in definitive proxy or information statements incorporated by reference in Part III of this Form 10-K or any amendment to this Form 10-K.

Indicate by check mark whether the registrant is a large accelerated filer, an accelerated filer, a non-accelerated filer, a smaller reporting company or an emerging growth company. See the definitions of "large accelerated filer," "accelerated filer," "smaller reporting company," and "emerging growth company" in Rule 12b-2 of the Exchange Act.

Large accelerated filer

Accelerated filer

Non-accelerated filer

Smaller reporting company

Emerging growth company

If an emerging growth company, indicate by check mark if the registrant has elected not to use the extended transition period for complying with any new or revised financial accounting standards provided pursuant to Section 13(a) of the Exchange Act.

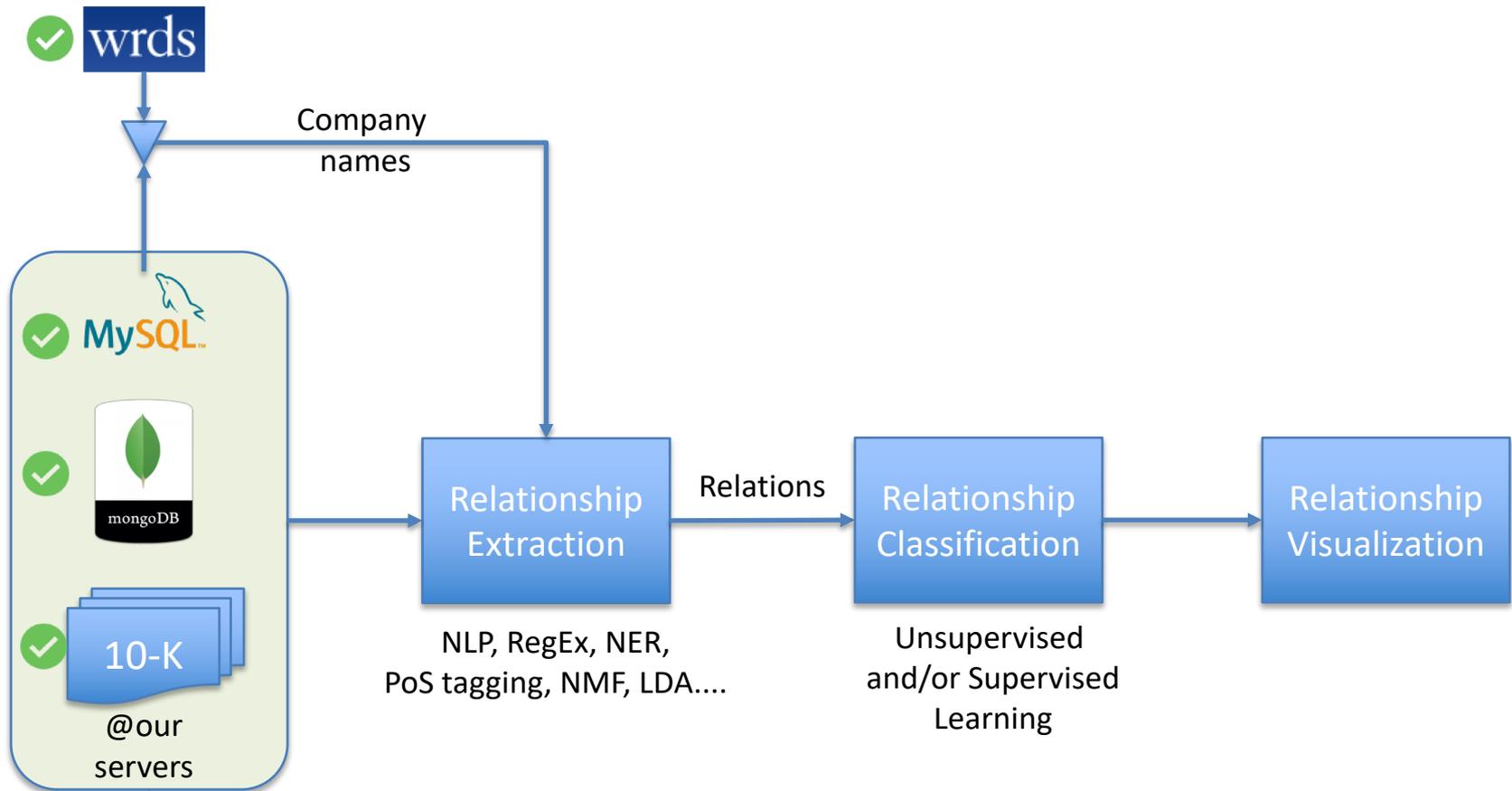
Indicate by check mark whether the registrant is a shell company (as defined in Rule 12b-2 of the Exchange Act). Yes No

As of June 29, 2018, the aggregate market value of shares held by non-affiliates of the registrant (based upon the closing sale prices of such shares on the Nasdaq Global Select Market on June 29, 2018) was approximately \$680.0 billion. For purposes of calculating the aggregate market value of shares held by non-affiliates, we have assumed that all outstanding shares are held by non-affiliates, except for shares held by each of our executive officers, directors and 5% or greater stockholders. In the case of 5% or greater stockholders, we have not deemed such stockholders to be affiliates unless there are facts and circumstances which would indicate that such stockholders exercise any control over our company, or unless they hold 10% or more of our outstanding common stock. These assumptions should not be deemed to constitute an admission that all executive officers, directors and 5% or greater stockholders are, in fact, affiliates of our company, or that there are not other persons who may be deemed to be affiliates of our company. Further information concerning shareholdings of our officers, directors and principal stockholders is included or incorporated by reference in Part III, Item 12 of this Annual Report on Form 10-K.

As of January 31, 2019, there were 299,360,029 shares of the registrant's Class A common stock outstanding, 46,535,019 shares of the registrant's Class B common stock outstanding, and 349,291,348 shares of the registrant's Class C capital stock outstanding.

DOCUMENTS INCORPORATED BY REFERENCE

Portions of the registrant's Proxy Statement for the 2019 Annual Meeting of Stockholders are incorporated herein by reference in Part III of this Annual Report on Form 10-K to the extent stated herein. Such proxy statement will be filed with the Securities and Exchange Commission within 120 days of the registrant's fiscal year ended December 31, 2018.



EDGAR

RELATIONSHIPS

Year	Cited_by	Citing	Text
2018	Alphabet	Microsoft	<p>Our business is characterized by rapid change as well as new and disruptive technologies. We face formidable competition in every aspect of our business, particularly from companies that seek to connect people with online information and provide them with relevant advertising. We face competition from:</p> <ul style="list-style-type: none"> • General purpose search engines and information services, such as Baidu, Microsoft's Bing, Naver, Seznam, Verizon's Yahoo, and Yandex. • Vertical search engines and e-commerce websites, such as Amazon and eBay (e-commerce), Booking's Kayak (travel queries), Microsoft's LinkedIn (job queries), and WebMD (health queries). Some users will navigate directly to such content, websites, and apps rather than go through Google.

RELATIONSHIPS

- **Ownerships**

In January 2018, we completed the acquisition of a team of engineers and a non-exclusive license of intellectual property from HTC for \$1.1 billion in cash. (Alphabet 2018)

- **Competition**

- **Agreements**

We also develop and manage our alliance relationships across a broad range of technology providers, including Amazon Web Services, Apple, Google, Microsoft, Oracle, Pegasystems, salesforce.com, SAP, Workday and many others (Accenture 2017)

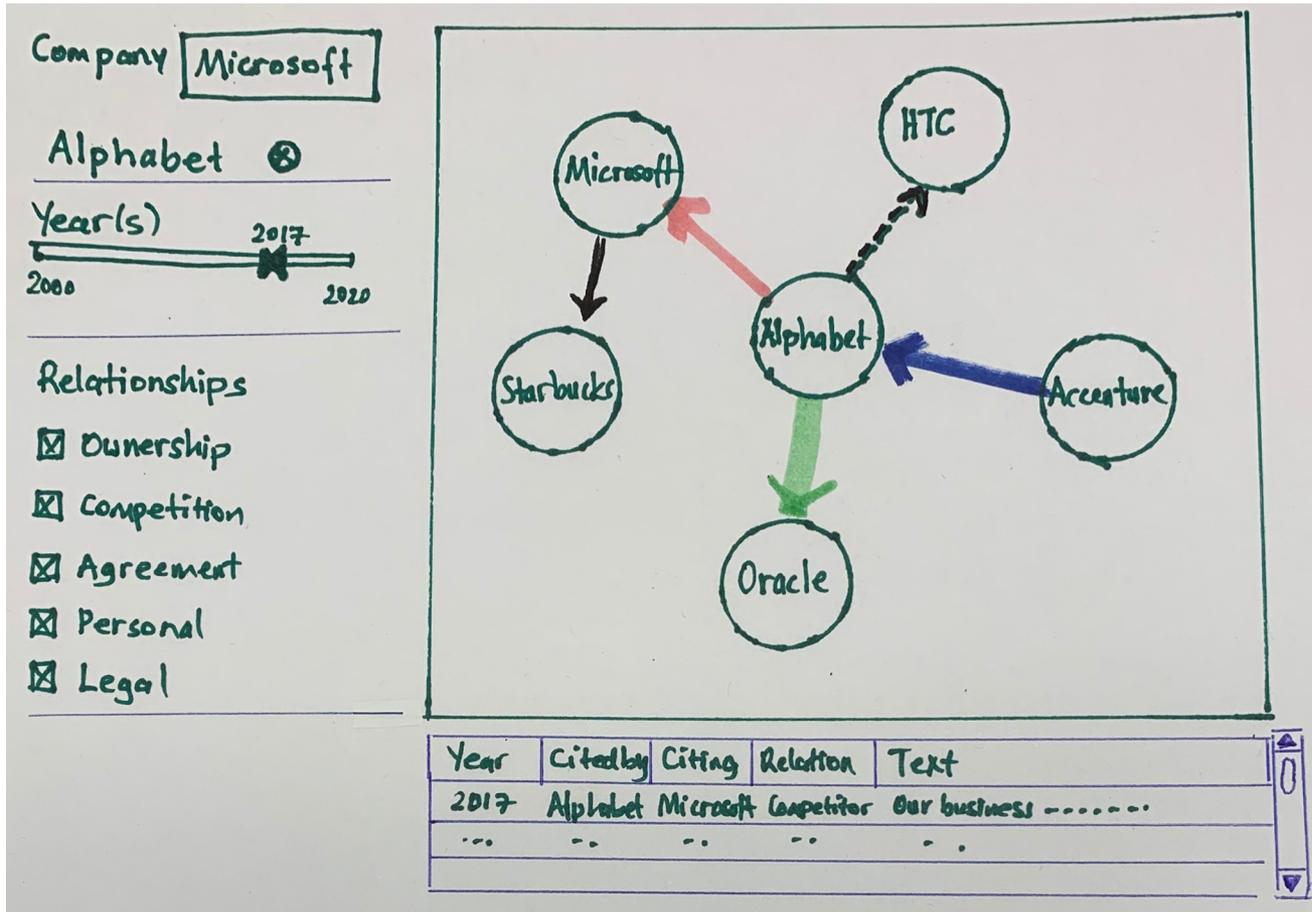
- **Personal Connections**

Mr. Nadella was appointed Chief Executive Officer in February 2014. Mr. Nadella also serves on the Board of Directors of Starbucks Corporation. (Microsoft 2017)

- **Legal Disagreements**

Oracle America, Inc. (Oracle) brought a copyright lawsuit against Google in the Northern District of California, alleging that ... (Aphabet 2018)

VISUALIZATION SKETCH



COMPUTING ENVIRONMENT

- [server_1_name].ubc.ca
 - Debian Linux server co-located at UBC Data Centre
 - Specs: 20 cores, 256 GB
 - SSH, Web, MySQL, MongoDB
- [server_2_name].ubc.ca
 - Debian Linux server co-located at UBC Data Centre
 - Specs: 16 cores, 128 GB, TITAN GPU
 - SSH, Web, MySQL
- Limited Access to Compute Canada or AWS for further computational needs if needs be
 - GPU

WHAT DO WE EXPECT

- Data products
 - Data pipeline
 - Data Science model(s) for relationship classification
 - Relationship Data
- Visualization
- Report

HOW WE WOULD HELP

- Domain Knowledge
- Supervision
- Process: Regular Meetings

ACKNOWLEDGEMENT

- Our research is generously funded by



QUESTIONS

- Thanks!
- Contact Information:
 - Hasan Cavusoglu: cavusoglu@sauder.ubc.ca
 - Gene Moo Lee: gene.lee@sauder.ubc.ca