

MACS 30122 PROJECT – FINAL PRESENTATION GROUP NAME: BOOM-SHAKA-LAKA PLEASE READ THE GLOSSARY IF YOU FIND ANY TERMINOLOGY CONFUSING

# STRUCTURE OF OUR REPORT

Background Introduction & Raw Data Collection

Barriers We Have Encountered

Processing Time, Legal Status, Inventor City, CPC

Network Analysis

Abstract Analysis (NLP)

User Interface



### BACKGROUND INTRO

- Patent analysis has been an important discipline in both managerial science and macroeconomics.
   Hence, we are interested in using patent data for technology analysis and planning and will explore in our project "Patent Stargazer" with a focus on cloud computation patents.
- We chose the top 3 companies with the largest market share in the cloud computation industry:







Amazon Technologies Inc.

Microsoft Technology Licensing, Llc

Google Inc.

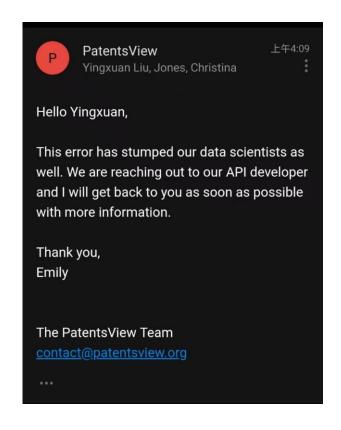
### RAW DATA COLLECTION

- This project has used 3 sources of data:
- The PatentsView API provides web developers and researchers programmatic access to longitudinal data and metadata on patents, inventors, companies, and geographic locations.
- incoPat provides as many as 260 searching fields, making an integration of data like patent litigation, licensing, etc.
- 1) incoPat Global Patent Database Manual download
- 2) PatentsView Patent API
- 3) PatentsView Assignee API





### BARRIERS WE HAVE ENCOUNTERED





While scraping for citation, the API imposed visit limitation and the service provider experienced internal issues which affects our schedule.

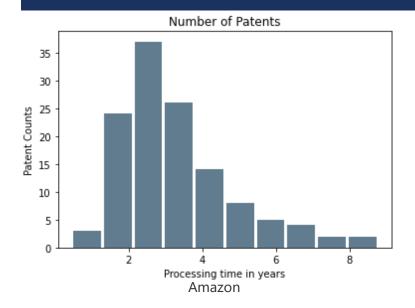


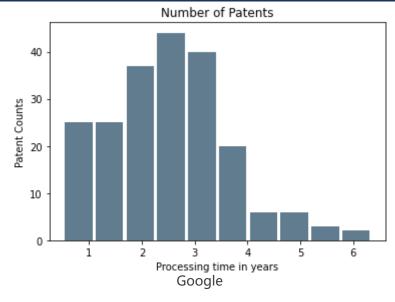
The packages for network analysis is slow while processing massive amount of data.

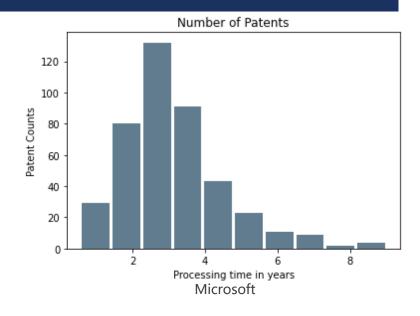


Nobody in the group had prior experience with GUI.

### PROCESSING TIME

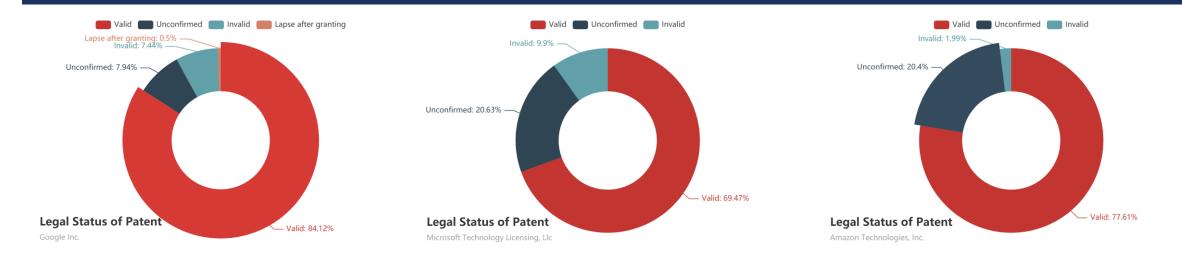






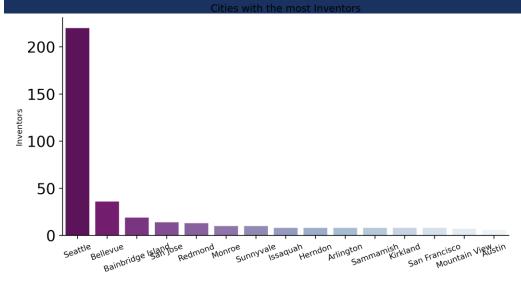
- How long are Amazon's patents usually processed?
- Mainly 2-3 yrs
- This time length is appropriate and shows that Amazon maintains a moderate speed with applying its patents rather than acquiring through litigation.

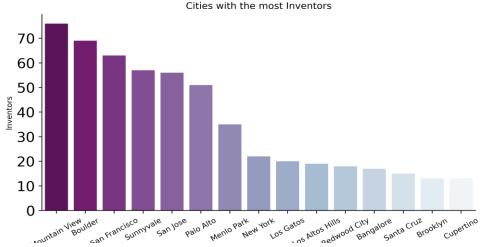
# LEGAL STATUS

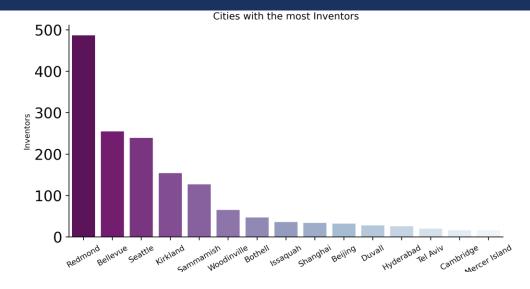


- Knowing the legal status actually explains some glitches happened during scraping.
- Having highest share of valid patents is good? (Like Google) Wait a sec...
- Drawn via Pyecharts

# **INVENTOR CITY**

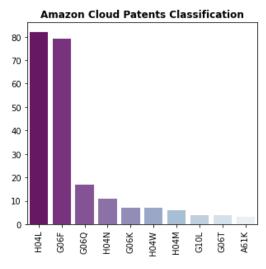


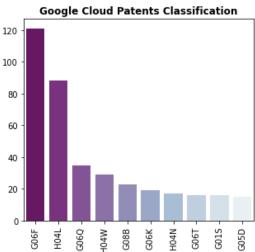


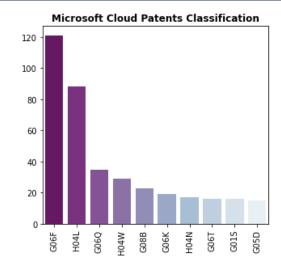


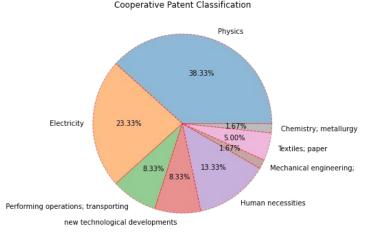
- Where are Amazon's cloud computation patents' inventors are at?
- Mainly the coastal cities
- Cities which are tech-heavy and have good tech and patent policies.
- \*Useful for industrial economics

# COOPERATIVE PATENT CLASSIFICATION (CPC)





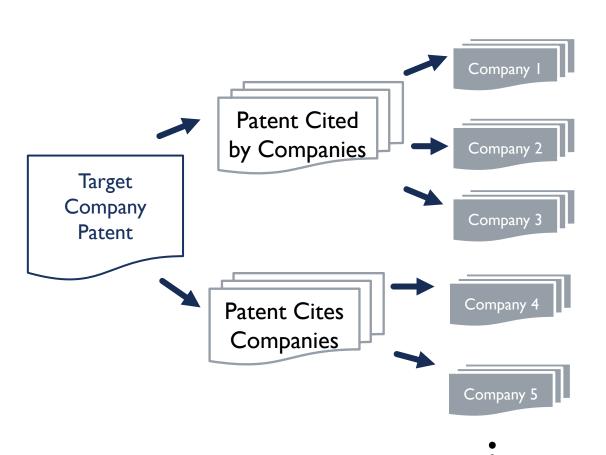




#### Method:

- According to the CPC, we visualize
   Amazon, Google and Microsoft 's CPC group id using PatentsView Patent API
- This is useful for determining the main application field of cloud computing
- Identify the diversified competition strategies for each targeted company
- CPC subclassification and section

### **NETWORK ANALYSIS**



#### **Data Collection:**

- Snowball sampling + UPSTO API
- Around 8,000 pairs of citation pairs
- 200 companies

#### **Data Processing:**

Networkx and Gephi

#### **Data analysis:**

- degree distribution
- centrality metrics
- smallworld coefficient
- community detection Modularity algorithm

$$Q = rac{1}{4m} \sum_{vw} B_{vw} s_v s_w = rac{1}{4m} \mathbf{s}^{ ext{T}} \mathbf{B} \mathbf{s},$$

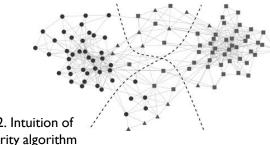


Figure 2. Intuition of modularity algorithm

Figure 1. Snowball sampling

Johnson Controls Taxanolesin Figure Victorial view in Equit. V. -

Think Company Corporation -Protegy Superior System Inc. ATLASSIA PTYLTD Carricol**®** Corp. -Сеогдержанд Оптива в Ирково в приняти на надвиси со приняти н United States of America as Represented by the Administrator of Hattanal Accountation and Space Administration Zecal@ Inc. -Haura (8, 160MY CORPORATION / PANZUES, 190, . Viewsoins Corporation -Voltage Salpety for . Ball Products, L.P. -.9 Fearing Committee Co. Line Investigate Inc. Today (149), key / Market World Trody (1) Northern Tallican Limites -Capette Buill Charact be Bostonel Austrastal Technicketheries Jrc. Koure no. Haderel yelgindre Corporation ... Katalania Ka disting the sensity of Africa . Nortel HetBorks S.A. -- ATKI Na 🏶 (a ÎST C.) Juriper Harmonday 2003 F110 -QUANTA COMPUTER INC. -Stebel Systems, Inc. -INSIDE & DURE -200Minio Internica. -Med HR. 13 Pears regarding 111 Several Research To Board of Torrigors of the district Board Board unique being that VATSUBHITA HILLSTRY COR-The Trustees of the Un Straity of Pennsylvania - , (Schilling - ) ( the transmitted by a series and the series of the series ALCOTEL : 665.84**8**0.110 -Helica College Ltd. -The Asicspace Corporation - Company Cable Communications McC -Do note accompanies touristicourisms in responsibles, lifet corporador viscourisms, he JUPITER NETWORKS, 190. снеі∰ічс. .. SEPARON. NO Extractor in a second of the contractor of the c Lega<mark>(0</mark>, 40) IVPRIVATA-INC. --Symantec Corporation covacci e collegia Salmirano di raprati on -Guill-Mileigenarion Systems inc. DatCard Sellierra, Inc. -SOM COMPLITER DIFFERTAINMENTING. Shap Laborderic of America, no. ENCIP HEREN CONSTRUCTOR PROCEDURED TO THE PROPERTY OF THE PROP PrivacyCeles@stern U.C.... Aer, Pre. TeleChern International Inc. You be inc. Spoul Apuillien, Inc. -Criscia serificite, force Dietaki Sel-Bereke Sygona Olika on al Diebold, recoporalet -East of the and Telecommunities of a Research Instance of Companion (1994) Approcked to the Si kan Constitution, he - trager of the key groups Pain Veckel Sautions, LLC San Vicrosprage Inc. Korec Institute of Science and Technology Information -Heave Indon'Specific, 118. PANASCHIS AMBHIBE CONTONATION Making At i 60 cmis Room and Laboratoring Los-RAF Techallogy, Inc. "BYTHE RED" MARTIN COMPANY LTD. Service Separation Companion. 💯 anglia naka 📆 digir mili n 🗇 Paleon Stor Setterna, Inc. Serretuce halfworks, Inc. Vertear Parant still Licensing Inc. Over Trailing on the SAPING Adobe Systems incorporated - Huel Us. Apple Comparering, Marcell Land RSA Spolling LLC ANY AND HOM DISENTENDED INC. HisrCia IV Profit-oldings, Inc. bint Ø lan Sun Victor<mark>e</mark> terre, inc. Way Committed are but Margaal, Standard on Served Instituted Congression . St CCK no His Sort Made Technolidad Established und Sort los Acpang Gistings Systems (\$250-50), his -Tergel B 🏶 du Inc. -Company to V PULITEO DI MITEO. Fordritt, Inc. -Aurigin Symeru, Inc. Larbes (Daffig), birtibeles Winnerson, ber-Managoret I de Bou Louis (M.S. Sylvanos, informació i una differencia Sylvanos i magnació a Lig. -Security 💏 at Corp. -Proof Sections Comparation Constitutes.

THE HEISEN SIGN THE UNIVERSITY OF MICHIGAN. HIGH SANGET CATAGORY. The C.R. Inc. Location Sering Corp. -Avid Tebrical Self Technique Research Foundation EMMANAGE OF THE PROPERTY OF THE POST OF TH Sdodffins -Branco de Aza**li**kurios, LLC -NORTH HICANADAS CORPORATION ... ISRUE, ACRESPACE MOUS THE SAIRS MAN. How But all States of Science of the Want of the Want of the Sair of Science of Science of the Sair of Science of Sc ATAT BEEL ECTUS PROPERTYLL P. hamacont and inca. Inc. -The Trace Brooks Along p. Inc. Charlestoner Internessing 100 and Berthart Surgery ATLAS ELEKTRONIK GMBII -Macc Mining Respect tion : Construit Residences, Inc. The Cooling Company -Sum of the House of the Party of the Conference of the Corporation -Many Harman, Harring Street, 18 Street, 18 Harring Cont. Space power Co. LTD -TOKYO ELECTRON LIMITED, the lad Storing of America at regressived by the Secretary of the Army -Baktoner Baltaner, Inc. The Same of Marin, William of Calenda, in normal Deception of the Calendary Products inc. Severi etterere -Surek Co. Inc. -MIZHE BEHAMBER AND CONTRACTOR STATE OF THE S MEC Compraign CI A perica -HON AS MOURE CO., 1 TO. Pay Executor Holistops (1955) Lie. - Manufacture Holistop (Laboratories of M. F. Cover (1964) United Technologies Corporation -Jail Propulator Laboratory -Missean Balladia Inn.

Signatus Plassey Cheffor Mayelerre El fileth 4 Nove. Simes' Test Group -

## **NETWORK ANALYSIS**

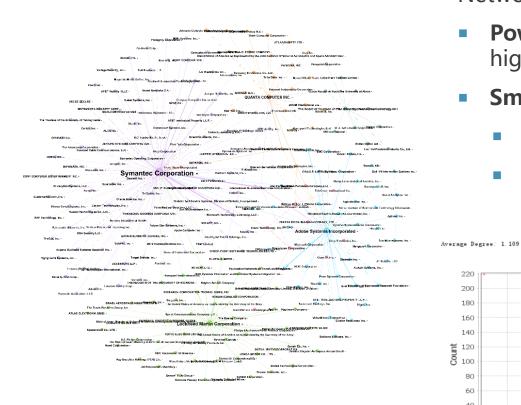


Figure 1. Company network

#### **Network Properties:**

- Power-law distribution of degrees: few companies with significantly high degree than average companies, forming hubs in networks
- Small-world effect:
  - Average length path is about one
  - check the small world robustness:  $_{\sigma}=1$  if  $_{\sigma}>1$  ( $C\gg C_{r}$  and  $Lpprox L_{r}$ ).

 $r = \frac{\frac{C}{C_r}}{L}$  small world holds true

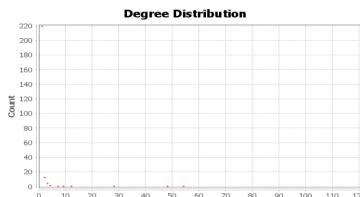


Table 1. Network Properties

	Cloud Computing Network	Random Network
Nodes / Edges	247 / 274	247 / 342
Average degree	1.109	1.385
Clustering Coefficient	0.005	0.001
Average path length	1.018	2.271
Network small- worldness	11.154	

# **NETWORK ANALYSIS**

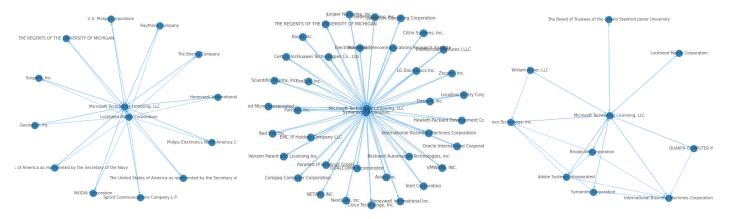


Figure I. Subgraph network( Left: Lockheed Martin; Center: Symantec; Right: Microsoft)

	company	pagerank score	company	betweeness centrality	company	eigenvector centrality	company	degree centrality
0	Microsoft Technology Licensing, LLC	0.442151	Microsoft Technology Licensing, LLC	0.827405	Microsoft Technology Licensing, LLC	0.827692	Microsoft Technology Licensing, LLC	1.657436
1	Cisco Technology, Inc.	0.000832	Symantec Corporation	0.068618	Cisco Technology, Inc.	0.415132	Symantec Corporation	0.126154
2	Sun Microsystems, Inc.	0.000788	Lockheed Martin Corporation	0.032503	International Business Machines Corporation	0.415132	Lockheed Martin Corporation	0.057436
3	Amazon Technologies, Inc.	0.000769	Adobe Systems Incorporated	0.016358	Oracle International Corporation	0.414874	Adobe Systems Incorporated	0.052308
4	International Business Machines Corporation	0.000762	QUANTA COMPUTER INC.	0.014871	Amazon Technologies, Inc.	0.414874	QUANTA COMPUTER INC.	0.029744
5	Broadcom Corporation	0.000738	The Board of Trustees of the Leland Stanford J	0.005099	Hewlett-Packard Development Company, L.P.	0.414874	The Board of Trustees of the Leland Stanford J	0.014359
6	Microsoft Corporation	0.000738	William Reber, LLC	0.002549	Microsoft Corporation	0.414617	Broadcom Corporation	0.011282
7	William Reber, LLC	0.000738	Broadcom Corporation	0.001700	Novell, Inc.	0.414617	William Reber, LLC	0.009231
8	OpenTV, Inc.	0.000726	Aptana, Inc.	0.001062	Intel Corporation	0.414617	Aptana, Inc.	0.006154
9	Canon Kabushiki Kaisha	0.000717	Alcatel USA Sourcing, L.P.	0.000000	Honeywell International Inc.	0.414617	Cisco Technology, Inc.	0.006154

Subgraph analysis and centrality measures:

- Top market shares =/= cuttingedged techniques.
  - Companies with highest market shares are not companies who have more citation and more innovation of cloud computing techniques.
  - It implies that middle and small size companies like Symantec are more expertized in cloud computing but its business is not competitive as the big-name companies like Amazon and Google.

# WORD CLOUD

#### Extract abstract and title

Focus on the cloud computing related abstracts

#### Text Normalization Pipeline

- Tokenization
- Lemmatization
- Remove stop words / punctuation
- POS: NN + VB

Visualization

#### Patents Raw Data



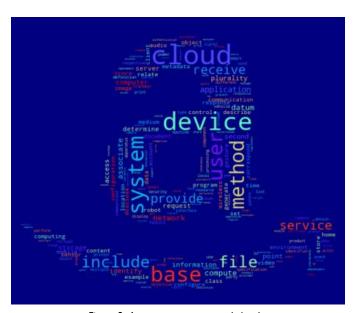


Figure 2. Amazon patents – word cloud



Word counts, distribution

Figure 3. Microsoft patents – word cloud

### USER INTERFACE & FUNCTIONAL DEMONSTRATION

- We build a visualization display interface with the GUI framework tkinter and PIL.
- Tkinter has several strengths. 1) It's crossplatform, so the same code works on Windows, macOS, and Linux. 2) It's lightweight and already built into the Python package.
- A main page where users can click on each button to quickly view each company and three sub-pages.
- Let's play!



### REFERENCES

- Mary Ellen Mogee (1991) Using Patent Data for Technology Analysis and Planning, Research-Technology Management, 34:4, 43-49, DOI: 10.1080/08956308.1991.11670755
- Muro, M., Maxim, R., & Whiton, J. (2019). Automation and artificial intelligence: How machines are affecting people and places.



# THANK YOU!

FEEL FREE TO REACH OUT TO ANY OF US IF THERE ARE QUESTIONS, THE GROUP MEMBERS ARE:

**CHUQING ZHAO** 

YI QIAN YINGXUAN LIU