

Prof. Dr. Alexander Hillert

Textual Analysis

Chapter 3: Database overview

July 18, 20, and 22, 2022

WHU – Otto Beisheim School of Management

Database overview

1. Newspapers, magazines, newswires
 - LexisNexis
 - Factiva
 - Proquest historical newspaper archive
2. EDGAR system of the SEC
 - Common types of filings
 - Linking CRSP/Compustat with EDGAR
 - File and folder structure
 - Identification of files
3. Other databases

1. Newspapers, magazines, newswires



LexisNexis, Factiva, Proquest

LexisNexis and Proquest should be accessible via the webpage of WHU's library
<https://www.whu.edu/de/programme/studentisches-leben/bibliothek/>

Einen Überblick über das Angebot erhalten Sie hier:

ABI/Inform Proquest	Internationale wirtschaftswissenschaftliche Medien
EBSCO Business Source	120 peer-reviewed Zeitschriften im Bereich Management und Wirtschaft
EIU Country Reports	EIU Länderberichte; politische und volkswirtschaftliche Analysen und Prognosen für 189 Länder
FT Financial Times	The Financial Times online (FT) ist eine internationale Tageszeitung für Wirtschaftsnachrichten
JSTOR I-VI	Zeitschriftenarchiv Packet "Art + Sciences"
Lexis Nexis	Wirtschaftsnachrichten, juristische Zeitungen und Nachrichtenmagazine, aus den USA und der ganzen Welt sowie lokale Publikationen

On the subsequent slides, you see screenshots using a Goethe Uni account which should, however, not result in major differences.

LexisNexis (1)

LexisNexis (now called Nexis)

- You (probably) need to be connected to WHU's network and/or login in with your WHU account to get access!

SUCHE NACH DATENBANKEN

Schnelle Suche

Go!

› [Erweiterte Suche](#)

› [Aktuelles](#)

› [Fachübersicht](#)

› [Alphabetische Liste](#)

› [Sammlungen](#)

› [Hinweise zur Benutzung](#)

› [Ansprechpartner](#)

› [Bibliotheksauswahl / Einstellungen](#)

› [Über DBIS](#)

Gefördert durch:



› [Impressum](#)

› [Datenschutz](#)

Ergebnisse Ihrer Suche

Sie suchten nach folgenden Kriterien:

- Suche über alle Felder: "nexus"

[Suche verändern](#)

Die Datenbank ist ...

- im Netz der Goethe-Universität zugänglich oder mit Login

Bitte beachten Sie die Nutzungsbedingungen und Copyright-Bestimmungen des Verlages/Herausgebers! +

Gewählte Datenbanken (2 Treffer)

	Zugang
Nexis Uni	Uninetz ●
Wolters Kluwer Online	Uninetz ●



LexisNexis (2)

Advanced Search

Home / Advanced Search
Advanced Search

All News Company and Financial Legal

Searching All

Microsoft

AND Enter your search term

Date Range

Date is after

Apr 01, 2020

Source

USA Today X The Washington Post X The New York Times X

Wall Street Journal Abstracts X

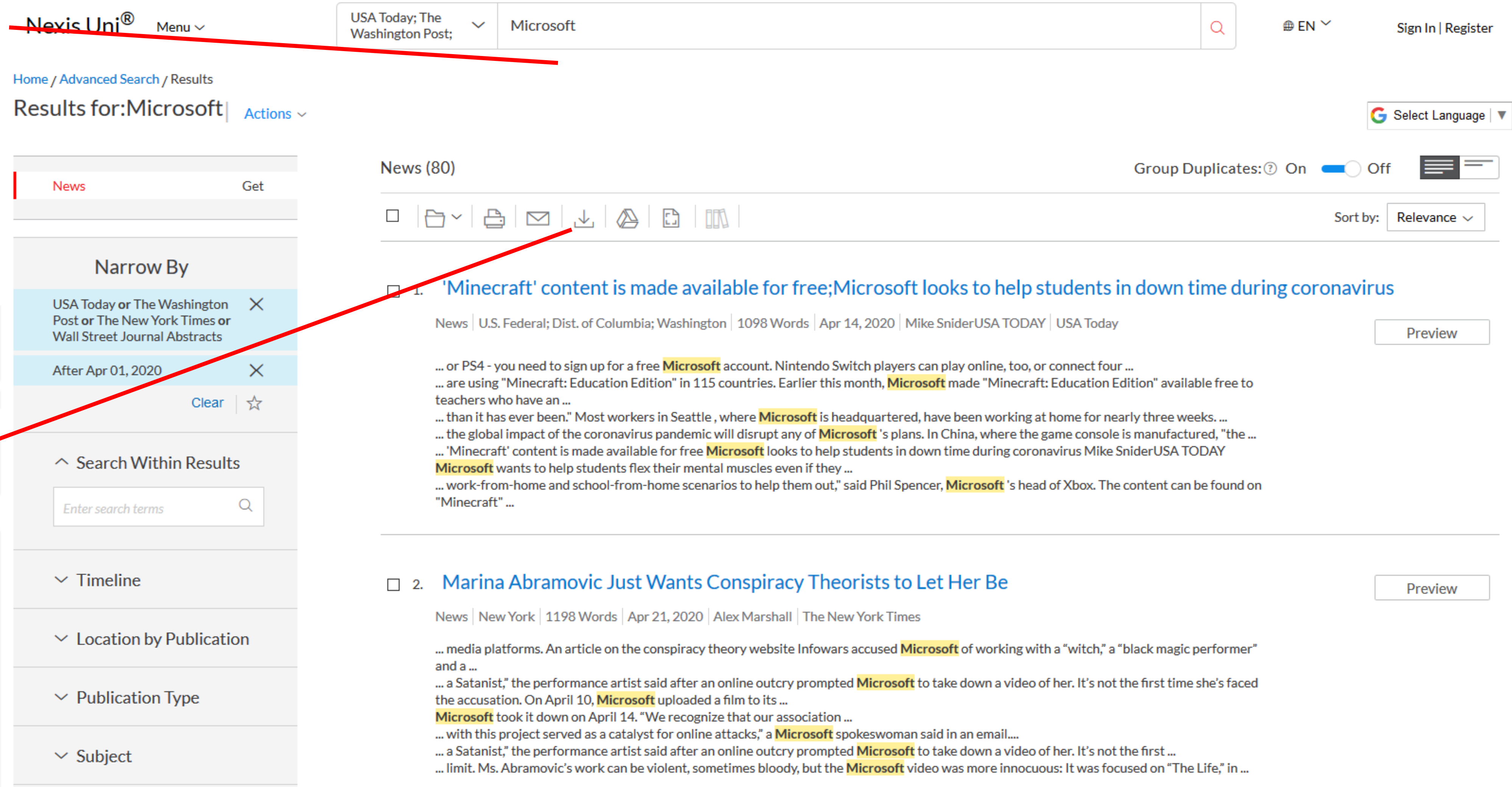
Enter one or more sources; then select from the list that appears

Enter search term

Select time period

Select sources

Modify search request



The screenshot shows the LexisNexis search interface. At the top, the search bar contains 'Microsoft' and the selected sources are 'USA Today; The Washington Post;'. A red box highlights the search bar area with the text 'Modify search request'. Below the search bar, the results are for 'Microsoft' with 80 news items. A red line points from the search bar to the first result. The first result is titled "'Minecraft' content is made available for free; Microsoft looks to help students in down time during coronavirus". Below the title, there is a 'Preview' button. The second result is titled 'Marina Abramovic Just Wants Conspiracy Theorists to Let Her Be' and also has a 'Preview' button. On the left side, there is a 'Narrow By' section with filters for 'USA Today or The Washington Post or The New York Times or Wall Street Journal Abstracts' and 'After Apr 01, 2020'. Below this is a 'Search Within Results' section with a search input field. At the bottom left, there are sections for 'Timeline', 'Location by Publication', 'Publication Type', and 'Subject'. A red box on the left side of the page contains the text 'Click here to download articles → new window will open', with a red line pointing to the download icon in the toolbar above the first result.

Click here to download articles → new window will open

LexisNexis (4)

Example

- Text in bold indicates searchable items.
- Percentages in parentheses (e.g., “Subject”, “Company”) indicate relevance scores.

20 stocks to buy in

USA TODAY

December 30, 2019 Monday, FIRST EDITION

Copyright 2019 Gannett Company, Inc. All Rights Reserved

Section: MONEY; Pg. 1B

Length: 1318 words

Byline: Jessica MentonUSA TODAY

Body

After a stellar 2019, investors look ahead to 2020 for stock picks. Analysts are skeptical that the stock market's gains, which are at more than 20% in 2019, will remain in the double-digit percentage range next year. They expect volatility to return in the midst of a U.S. presidential election. Some investors wait for stocks to get cheaper before they step in and scoop up buying opportunities. Companies poised to outperform, they say, will be ones that can continue to grow their earnings even if the economy slows. From iPhone maker Apple to beverage giant Coca-Cola to e-commerce titan Amazon, USA TODAY offers 20 stock picks for 2020 based on research reports and interviews with Wall Street stock analysts.

Apple (AAPL)

Strong demand for Apple's 5G iPhone next fall is projected to boost the company in 2020, says Piper Jaffray analyst Michael Olson. In December, Olson reiterated his "overweight" rating on the stock and raised his 12-month price target to \$305 from \$290, citing robust iPhone demand and better-than-expected performance in wearables such as the Apple Watch, AirPods and AirPods Pro. Apple shares finished Thursday at \$289.91.

Classification

Language: ENGLISH

Publication-Type: NEWSPAPER

Subject: SECURITIES & OTHER INVESTMENTS (91%); INTERNET RETAILING (90%); MANUFACTURING OUTPUT (90%); REPORTS, REVIEWS & SECTIONS (90%); COMPUTER & SOFTWARE STORES (90%); STOCK EXCHANGES (90%); BROKER RECOMMENDATIONS (90%); INVESTMENT ADVISERS (90%); INDUSTRY ANALYSTS (89%); BUSINESS FORECASTS (89%); STOCK PRICES (89%); HOLDING COMPANIES (89%); NEW ISSUES (89%); SECURITIES BROKERS (88%); FINANCIAL RATINGS (78%); DIVIDENDS (78%); OUTPUT & DEMAND (78%); COMPANY EARNINGS (78%); INVESTMENT MANAGEMENT (78%); SALES FIGURES (78%); SALES & SELLING (77%); US PRESIDENTS (74%); RESEARCH REPORTS (73%); ECONOMIC DECLINE (70%); BUSINESS REPORTS & FORECASTS (58%); US PRESIDENTIAL ELECTIONS (57%); ELECTRONIC COMMERCE (55%)

Company: APPLE INC (96%); AMAZON.COM INC (84%); COCA-COLA CO (84%); MICROSOFT CORP (82%); WALT DISNEY CO (81%); NIKE INC (80%); PIPER JAFFRAY COS (56%); BANK OF AMERICA CORP (55%)

Ticker: AAPL (NASDAQ) (96%); AMZN (NASDAQ) (84%); KO (NYSE) (84%); MSFT (NASDAQ) (82%); DIS (NYSE) (81%); NKE (NYSE) (80%); PJC (NYSE) (56%); BAC (NYSE) (55%); BAC (LSE) (55%); 8648 (TSE) (55%)

Industry: NAICS423430 COMPUTER & COMPUTER PERIPHERAL EQUIPMENT & SOFTWARE MERCHANT WHOLESALE (96%); NAICS334413 SEMICONDUCTOR & RELATED DEVICE MANUFACTURING (96%); NAICS334112 COMPUTER STORAGE DEVICE MANUFACTURING (96%); NAICS334111 ELECTRONIC COMPUTER MANUFACTURING (96%); SIC5045 COMPUTERS & COMPUTER PERIPHERAL EQUIPMENT & SOFTWARE (96%); SIC3674 SEMICONDUCTORS & RELATED DEVICES (96%); SIC3577 COMPUTER PERIPHERAL EQUIPMENT, NEC (96%); SIC3572 COMPUTER STORAGE DEVICES (96%); SIC3571 ELECTRONIC COMPUTERS (96%); NAICS454110 ELECTRONIC SHOPPING AND MAIL-ORDER HOUSES (84%); SIC5961 CATALOG & MAIL-ORDER HOUSES (84%); NAICS312111 SOFT DRINK MANUFACTURING (84%); SIC2086 BOTTLED & CANNED SOFT DRINKS & CARBONATED WATER (84%); NAICS511210 SOFTWARE PUBLISHERS (82%); SIC7372 PREPACKAGED SOFTWARE (82%); NAICS713110

Search terms

- ‘Microsoft’ → all articles containing the word ‘Microsoft’
- ‘Company(“Microsoft Corp”)’ → all articles assigned to Microsoft Corp.
- ‘Company(“Microsoft Corp (9*%)”)’ → all articles assigned to Microsoft Corp with a relevance score of 90 and more.
→ *As of April 23, 2021 this request seems no longer to work.*
- Always check whether the articles really match the company.
Example: Be Inc (permno 87101, gvkey 122075)
- You can use additional search filters, e.g., document length.

LexisNexis (6)

Herunterladen

Basis-Optionen Layout-Optionen

Was möchten Sie herunterladen?

Vollständige Dokumente
 Dokumentanhänge einschließen (wenn verfügbar)

1-100
(z. B. 1-25, 70, 245-300 [Weitere Informationen](#))

Hinweis: Sie haben die maximale Anzahl Dokumente 100 erreicht, die Sie auswählen können.

Nur Anhänge mit vollständigen Dokumenten (wenn verfügbar)
 Titelliste für 'News':
 Bibliografie einschließen

Dateityp

PDF
100 document limit.

MS Word (docx)
100 document limit.

Rich Text Format (rtf)
100 document limit.

Beim Herunterladen mehrerer Dokumente

Dokumente zusammenfassen und als einziges Dokument exportieren (Hinweis: Anhänge werden als separate Dokumente)

Die Verbreitung unterliegt den [Allgemeinen Geschäftsbedingungen](#)

Herunterladen Abbrechen

The maximum # of documents per download is 100 → split in batches of 100, i.e. 1-100, 101-200, ...

Herunterladen

Basis-Optionen **Layout-Optionen**

Hinweis: Anhänge können nur ein Deckblatt und den Vor- und Nachnamen enthalten. Ergebnislisten können nur das Deckblatt, den Vor- und Nachnamen, das Nexis Uni®-Logo und die Seitennummerierung im Dokument enthalten.

Folgende Elemente in gesendetes Dokument einschließen:

Deckblatt
 Vor- und Nachname auf Deckblatt und in Fußzeile
 Benutzernamen und Daten der Suche einschließen
 Seitennummerierung im Dokument

Oben rechts
 Unten in der Mitte
 Unten rechts

Eingebettete Referenzen als Links
 Eingebettete juristische Zitationen kursiv anzeigen

Schriftart

Arial

Schriftgröße

10

Zeilenabstand

Einfacher Zeilenabstand
 Doppelter Zeilenabstand

Layout

Suchbegriffe fett anzeigen
 Suchbegriffe kursiv anzeigen
 Suchbegriffe unterstreichen
 Seitennummern in Jurisprudenzsammlungen fett drucken

Die Verbreitung unterliegt den [Allgemeinen Geschäftsbedingungen](#)

Herunterladen Abbrechen

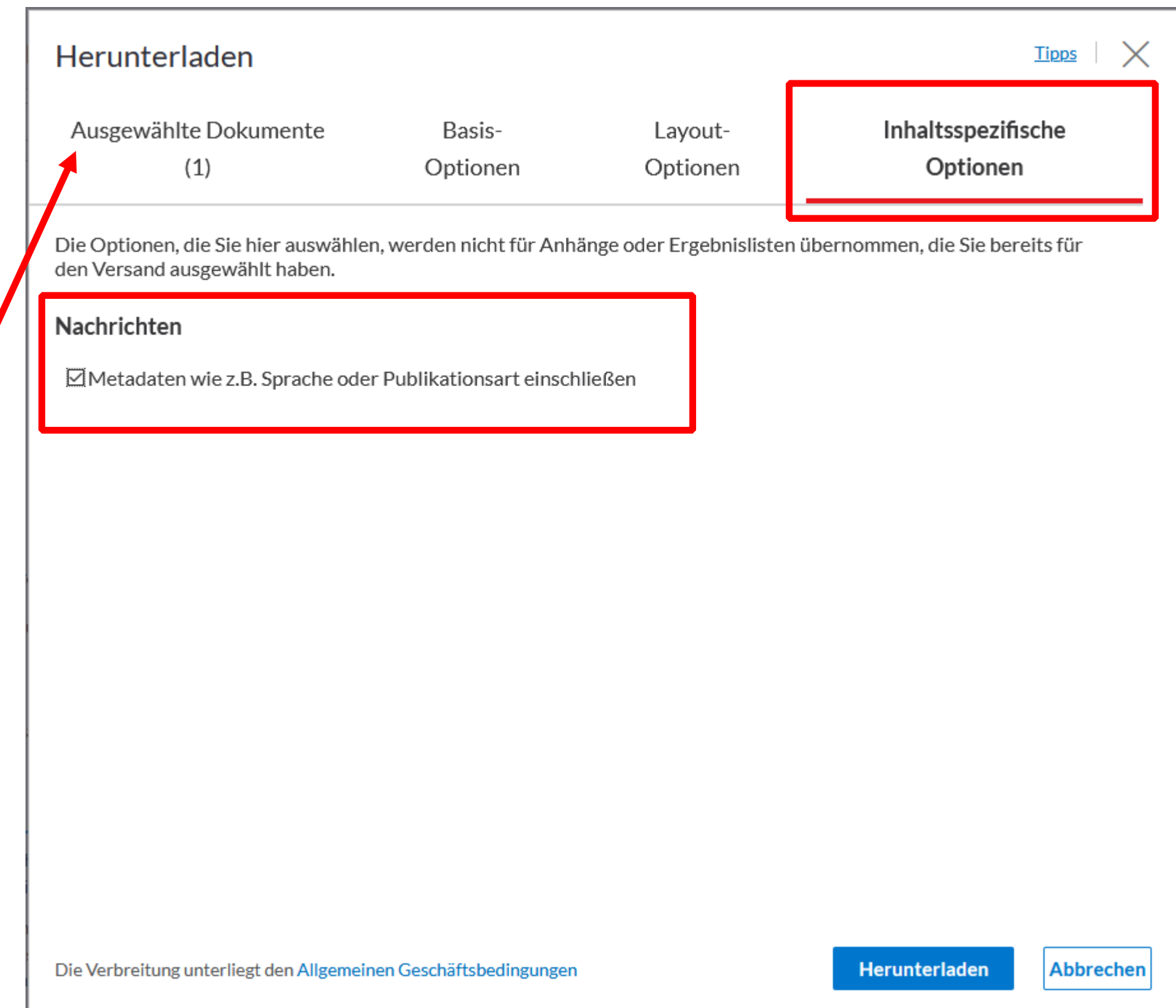
Use docx format → in Python

1. Install the docx2txt package:
`pip install docx2txt`
2. To open the docx in Python:
 - I. `import docx2txt`
 - II. `document_text = docx2txt.process("test.docx")`
 - III. `print(document_text)`

You can keep the default layout options.

Current issues with LexisNexis

- Database is slow to process requests.
- There seems to be no way to include the document details (see [slide 8](#)) in the batch download.
 - alternative: use individual downloads or click through the documents and copy-paste the information.



Herunterladen Tipps | ×

Ausgewählte Dokumente (1) Basis-Optionen Layout-Optionen **Inhaltsspezifische Optionen**

Die Optionen, die Sie hier auswählen, werden nicht für Anhänge oder Ergebnislisten übernommen, die Sie bereits für den Versand ausgewählt haben.

Nachrichten

Metadaten wie z.B. Sprache oder Publikationsart einschließen

Die Verbreitung unterliegt den [Allgemeinen Geschäftsbedingungen](#)

Herunterladen **Abbrechen**

Factiva

- Alternative data source for newspapers, newswires, and magazines.
- Coverage/sources of Lexis and Factiva are different.
- Factiva is not available at Goethe University!

The screenshot displays the Factiva search interface within a Mozilla Firefox browser. The browser's address bar shows the URL: `https://global.factiva.com/ha/default.aspx#.../!7&_suid=1431692203370008147640000847955`. The page header features the 'DOW JONES' logo and the 'FACTIVA' logo, along with navigation tabs for 'Home', 'Suche', 'Alerts', 'Nachrichtenseiten', and 'Unternehmen/Märkte'. The search bar contains the text 'abreast of the market' and shows search filters for 'DATUM: 01/01/2004 bis 31/01/2004' and 'QUELLE: The Wall Street Journal'. Below the search bar, there are options for 'Freitextsuche', 'Suchformular', and 'Beispiele'. The search results section shows 'Quelle: The Wall Street Journal' and a list of source categories: 'Alle Quellen', 'Alle Publikationen', 'Alle Webnachrichten', 'Alle Blogs', 'Alle Bilder', and 'Alle Multimedia-Nachrichten'. The bottom of the page shows a taskbar with various applications open, including 'Posteingang in staff.mail...', 'Factiva - Mozilla Fire...', 'Unbenannt - Paint', and 'Dokument1 - Microsoft ...'.

Factiva (2)

Back to search interface

To download the articles, you need to set the checkmarks first. The max. # is restricted to 25 or 50.

Comparison between LexisNexis and Factiva

LexisNexis

- Advantage
 - Relevance Score
 - Large number of data sources
 - ~~Easy to handle~~
 - ~~Allows for reasonable download volume~~
- Disadvantage
 - No full text of WSJ and no Dow Jones newswire.
 - Download of document details cumbersome.

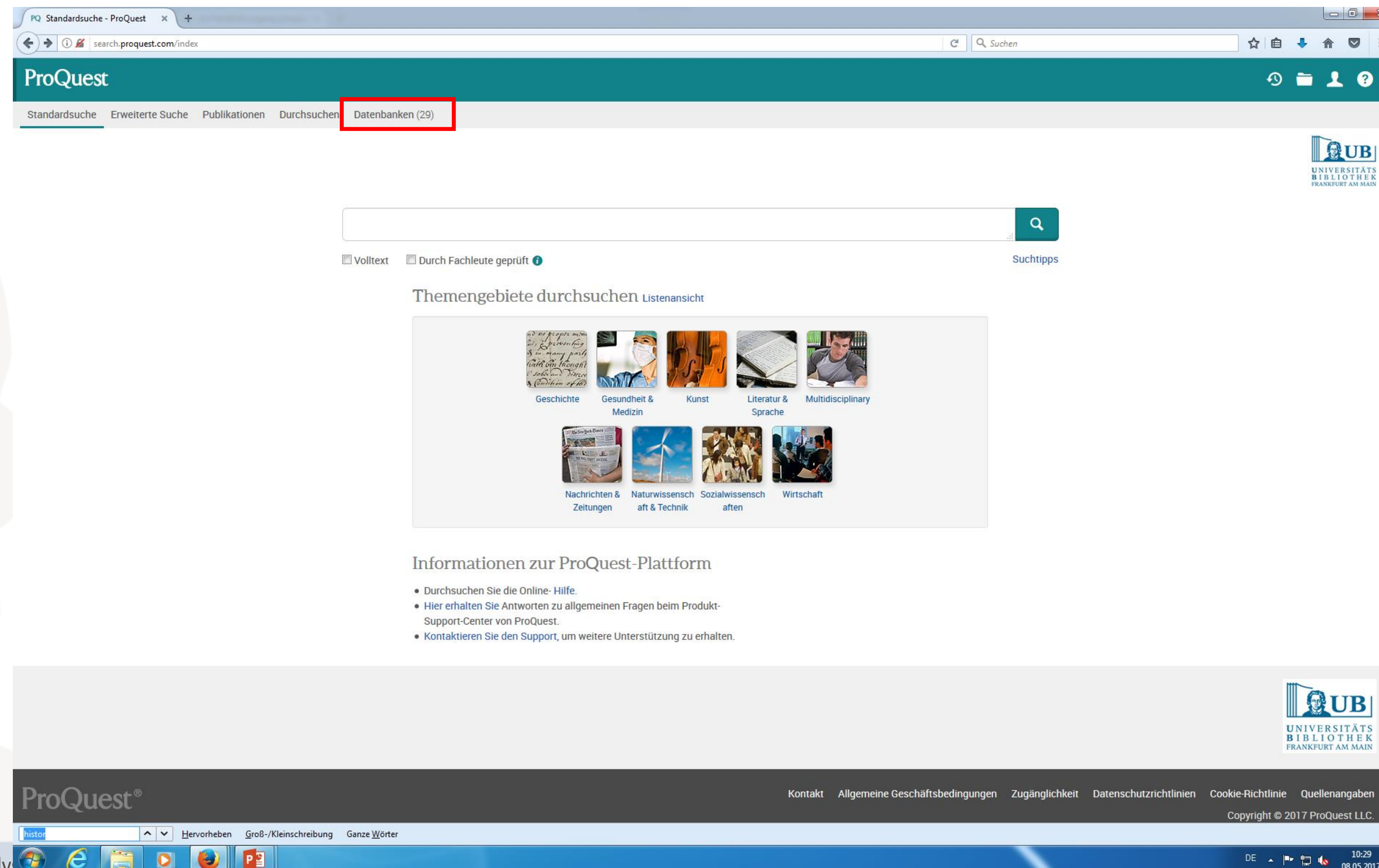
Factiva

- Advantage
 - Full text access to WSJ and Dow Jones newswire.
- Disadvantage
 - Very restrictive regarding download volume.

Proquest historical newspapers (1)

Proquest historical newspapers

- Accessible via WHU's library webpage or at <http://search.proquest.com/index>.
- "Datenbanken" → "ProQuest Historical Newspaper: XXX", e.g., "ProQuest Historical Newspaper: The New York Times"



Proquest historical newspapers (2)

Themengebiet(e): Sozialwissenschaften , Gesundheit & Medizin	
<input checked="" type="checkbox"/> ProQuest Dissertations & Theses Global ⓘ	
Multidisciplinary - dissertations Themengebiet(e): Multidisciplinary	
<input checked="" type="checkbox"/> ProQuest Historical Newspapers: The Boston Globe (1872 - 1985) ⓘ	
Historische Nachrichten - Zeitungsartikel Themengebiet(e): Geschichte , Nachrichten & Zeitungen Titelliste anzeigen	
<input checked="" type="checkbox"/> ProQuest Historical Newspapers: The Boston Globe (1872 - 1922) ⓘ	
Historische Nachrichten - Zeitungsartikel Themengebiet(e): Geschichte , Nachrichten & Zeitungen Titelliste anzeigen	
<input checked="" type="checkbox"/> ProQuest Historical Newspapers: Chicago Tribune (1849 - 1922) ⓘ	
Historische Nachrichten - Zeitungsartikel Themengebiet(e): Geschichte , Nachrichten & Zeitungen Titelliste anzeigen	
<input checked="" type="checkbox"/> ProQuest Historical Newspapers: The Guardian and The Observer (1791 - 1900) ⓘ	
Historische Nachrichten - Zeitungsartikel Themengebiet(e): Geschichte , Nachrichten & Zeitungen Titelliste anzeigen	
<input checked="" type="checkbox"/> ProQuest Historical Newspapers: Los Angeles Times (1881 - 1922) ⓘ	
Historische Nachrichten - Zeitungsartikel Themengebiet(e): Geschichte , Nachrichten & Zeitungen Titelliste anzeigen	
<input checked="" type="checkbox"/> ProQuest Historical Newspapers: The New York Times (1851 - 1922) ⓘ	
Historische Nachrichten - Zeitungsartikel Themengebiet(e): Geschichte , Nachrichten & Zeitungen Titelliste anzeigen	
<input checked="" type="checkbox"/> ProQuest Historical Newspapers: The Wall Street Journal (1889 - 1922) ⓘ	
Historische Nachrichten - Zeitungsartikel Themengebiet(e): Geschichte , Nachrichten & Zeitungen , Wirtschaft Titelliste anzeigen	
<input checked="" type="checkbox"/> ProQuest Historical Newspapers: The Washington Post (1877 - 1922) ⓘ	
Historische Nachrichten - Zeitungsartikel Themengebiet(e): Geschichte , Nachrichten & Zeitungen Titelliste anzeigen	
<input checked="" type="checkbox"/> Social Services Abstracts (1979 - aktuell) ⓘ	
Social work and human services - journal articles, dissertations, etc.	

Data availability until 1922 → probably hard to get capital market data.

Proquest historical newspapers (3)

Example: General Motors

Q

Volltext Durch Fachleute geprüft i
Suche ändern Vorherige Suchen Suche/Alert speichern ▼

Verwandte Suchen General Motors Corp (Company/Org) General Motors Corp (Company/Org) AND automobile industry General Motors Corp (Company/Org) AND Ford Motor Co (Company/Org)
Alle anzeigen >

739.198 Ergebnisse 18.727 ebrary-E-Books

Suchen in
Zitieren E-Mail Drucken Speichern

Relevanz

Sortieren

1-20 auswählen 0 ausgewählte Einträge

Kurzansicht |
 Detailansicht

Eingegrenzte Ergebnisse

Volltext ▼

Durch Fachleute geprüft ▼

Quellentyp ^

- Dissertationen & Abschlussarbeiten (390.878)
- Historische Zeitungen (230.768)
- Fachzeitschriften (89.835)
- Magazine (11.202)
- Historische Zeitschriften (10.913)
- Weitere Optionen...

Publikationsdatum ^

1693 - 2017 (Dekaden)

<input type="checkbox"/>	1	<div style="border: 1px solid #ccc; padding: 2px; margin-bottom: 5px;"> CHEVROLET MOTOR'S SHARE OF GENERAL MOTORS STOCK: CHEVROLET CO. IS UNDERSTOOD TO HOLD ABOUT \$8,000 SHARES OF GENERAL MOTORS COMMON STOCK. Of These 10,000 Shares Weree Acquired on Baris of Four- Chevrolet for One General Motors, and Balance, Five to One--Chevolst on-General Motors Selling for \$140 a Share--Big Demand for Cars </div> <div style="font-size: 0.8em;"> Wall Street Journal (1889-1922); New York, N.Y. [New York, N.Y.]19 Sep 1916: 2. ...General Motors recapitalization is the Chevrolet Motor Co. which for the past ...for 1 General Motors. </div> <div style="display: flex; justify-content: space-between; font-size: 0.8em;"> Kurzfassung/Details PDF Volltext - PDF (44 KB) Vorschau ▼ </div>
<input type="checkbox"/>	2	<div style="border: 1px solid #ccc; padding: 2px; margin-bottom: 5px;"> GENERAL MOTORS BUILDING: Name of Durant Building, Home Office of General Motors to Be Changed to General Motors Building </div> <div style="font-size: 0.8em;"> Wall Street Journal (1889-1922); New York, N.Y. [New York, N.Y.]25 Feb 1921: 5. ...of General Motors Corp. in Detroit, is to be changed to General Motors Building. </div> <div style="display: flex; justify-content: space-between; font-size: 0.8em;"> Kurzfassung/Details PDF Volltext - PDF (10 KB) Vorschau ▼ </div>
<input type="checkbox"/>	3	<div style="border: 1px solid #ccc; padding: 2px; margin-bottom: 5px;"> CHEVROLET'S AQUISION OF CONTROL IN GENERAL MOTORS: SLIGHTLY OVER 100,000 SHARES OF GENERAL MOTORS COMMON WILL BE ABSORBED AT THIS TIME Offer May Be Extended to All General Motors Common Stockholders to Exchange at 1 to 5 Ratio Not Many Preferred Shares to Be Acquired General Motors Earning \$125 a Share, Chevrolet \$25 </div> <div style="font-size: 0.8em;"> Wall Street Journal (1889-1922); New York, N.Y. [New York, N.Y.]23 Dec 1915: 7. ...General Motors by Chevrolet Motor Co. will involve an increase in the capital of </div> <div style="display: flex; justify-content: space-between; font-size: 0.8em;"> Kurzfassung/Details PDF Volltext - PDF (68 KB) Vorschau ▼ </div>

Proquest historical newspapers (4)

Newspaper articles are only available in pdf-format

- Need to be converted in txt format.
- Scanned pdf → optical character recognition software (OCR) needed, e.g., ABBYY.
- Depending on scan quality errors are likely (see discussion in García (2013)).



2. EDGAR system of the SEC



EDGAR system of the SEC (1)

Go to <https://sec.gov>

Click 'Company Filings'
and then "More
Options"

U.S. SECURITIES AND EXCHANGE COMMISSION

Search SEC.gov

COMPANY FILINGS

ABOUT | DIVISIONS & OFFICES | ENFORCEMENT | REGULATION | EDUCATION | **FILINGS** | NEWS

Upcoming Webcast
Open Meeting | Mar 21, 11:00 am EDT

SECURING FUTURES

The SEC enforces the securities laws to protect the more than 66 million American households that have turned to the securities markets to invest in their futures—whether it's starting a family, sending kids to college, saving for retirement or attaining other financial goals.

Information on the
filings 'Filings' →
'Description of SEC
forms' and/or 'SEC
Forms List (PDF
versions)

Overview of common forms (1)

Where to get information on the different forms?

- <http://www.sec.gov/info/edgar/forms/edgform.pdf> lists different forms with brief description.

Commonly used forms

- Annual and quarterly company reports (form 10-K and 10-Q filings)
 - Include also 10-K405 filings!
'Until this distinction was eliminated in 2003, a substantial portion of 10-Ks were categorized as 10-K405. The SEC eliminated the 405 classification due to confusion and inconsistency in its application.' Loughran and McDonald (2011, footnote 4)
 - There are simplified reports for small businesses (10KSB and 10QSB)

Overview of common forms (2)

Commonly used forms - continued

- Insider trading (form 3, 4, and 5 filings)
 - Standardized form
 - Default option: excluded from search
 - Available in the Thomson Reuters Insiders database (currently not subscribed)
- Institutional holdings (form 13-f filings)
 - Standardized forms
 - Example: Windsor Financial Group LLC <http://www.sec.gov/cgi-bin/browse-edgar?company=&match=contains&CIK=1132356&filenum=&State=&Country=&SIC=&owner=exclude&Find=Find+Companies&action=getcompany>

Overview of common forms (3)

Commonly used forms - continued

- Investment management companies
 - Portfolio holdings (form N-Q filings)
Available in the Thomson Reuters Mutual Fund Holdings (s12) - (currently not subscribed)
 - Annual and semi-annual shareholder reports (form N-CSR and N-CSRS)
See Hillert, Niessen-Ruenzi, and Ruenzi (2021) for details.
 - Annual and semi-annual report (form NSAR-B and NSAR-A)
 - Clear structure with numerical items
 - Example: Fidelity Magellan Fund <http://www.sec.gov/cgi-bin/browse-edgar?action=getcompany&CIK=0000061397&owner=exclude&count=40&hidefilings=0>

EDGAR system of the SEC (2)

Search interface

Select 'Contains' to not missing any filings

EDGAR | Company Filings

Company and Person Lookup

Name, ticker symbol, or CIK

Search Match Options

Starts with or Contains

File Number

To search by file number,
company name must be left
blank.

State

Country

Standard Industry Classification 

Ownership Forms 3, 4, and 5.

Include Exclude Only

Less Options ▲

SEARCH

How to Use this Search?

Enter name, ticker or CIK into the **single search field**.

Suggestions as you type link directly to filings.

EDGAR system of the SEC (3)

I recommend using the "Classic version"

Search result

FAQ | [EDGAR search tools](#) | [Feedback](#) | [Classic version](#)

| Company Search Results

[Home](#) » [Company Search](#)
MICROSOFT CORP MSFT on Nasdaq

Investor Toolkit: On ▼

[+] Company Information

Latest Filings (excluding insider transactions)

- February 10, 2022 - SC 13G/A: Statement of acquisition of beneficial ownership by individuals - *amendment* Filing
- February 8, 2022 - SC 13G/A: Statement of acquisition of beneficial ownership by individuals - *amendment* Filing
- January 25, 2022 - 10-Q: Quarterly report for quarter ending December 31, 2021 Filing
- January 25, 2022 - 8-K: Current report Filing ****Earnings release****
2.02 - Results of Operations and Financial Condition
9.01 - Financial Statements and Exhibits
- January 18, 2022 - 8-K: Current report Filing
7.01 - Regulation FD Disclosure
9.01 - Financial Statements and Exhibits

[View filings](#)

Selected Filings

- [+] 8-K (current reports)
- [+] 10-K (annual reports) and 10-Q (quarterly reports)
- [+] Proxy (annual meeting) and information statements
- [+] Ownership disclosures

EDGAR system of the SEC (4)

Search result – Classic version

SEC Home » Search the Next-Generation EDGAR System » Company Search » Current Page

MICROSOFT CORP CIK#: 0000789019 (see all company filings)

SIC: 7372 - SERVICES-PREPACKAGED SOFTWARE
 State location: WA | State of Inc.: WA | Fiscal Year End: 0630
 (Office of Technology)
 Get **insider transactions** for this issuer.
 Get **insider transactions** for this reporting owner.

Business Address
 ONE MICROSOFT WAY
 REDMOND WA 98052-6399
 425-882-8080

Mailing Address
 ONE MICROSOFT WAY
 REDMOND WA 98052-6399

Filter Results

Filing Type: Prior to: (YYYYMMDD) Ownership? include exclude only Limit Results Per Page

Search Within Files **EDGAR** | Full Text Search
 Enter keywords

Items 1 - 40 [RSS Feed](#)

Filings	Format	Description	Filed/Effective	File/Film Number
SC 13G/A	Documents	[Amend] Statement of acquisition of beneficial ownership by individuals Acc-no: 0001104659-22-018095 (34 Act) Size: 12 KB	2022-02-10	005-38758 22610358
SC 13G/A	Documents	[Amend] Statement of acquisition of beneficial ownership by individuals Acc-no: 0000834237-22-008433 (34 Act) Size: 13 KB	2022-02-08	005-38758 22602666
10-Q	Documents Interactive Data	Quarterly report [Sections 13 or 15(d)] Acc-no: 0001564590-22-002324 (34 Act) Size: 21 MB	2022-01-25	001-37845 22554019
8-K	Documents Interactive Data	Current report, items 2.02 and 9.01 Acc-no: 0001193125-22-017041 (34 Act) Size: 538 KB	2022-01-25	001-37845 22553943
8-K	Documents Interactive Data	Current report, items 7.01 and 9.01 Acc-no: 0001193125-22-010667 (34 Act) Size: 7 MB	2022-01-18	001-37845 22533739
S-8 POS	Documents	Securities to be offered to employees in employee benefit plans, post-effective amendments Acc-no: 0001193125-21-354016 (33 Act) Size: 65 KB	2021-12-10	333-185757 211485808
S-3ASR	Documents	Automatic shelf registration statement of securities of well-known seasoned issuers Acc-no: 0001193125-21-353912 (33 Act) Size: 598 KB	2021-12-10	333-261590 211485230
8-K	Documents Interactive Data	Current report, item 5.07 Acc-no: 0001193125-21-343711 (34 Act) Size: 276 KB	2021-11-30	001-37845 211459985
PX14A6G	Documents	Notice of exempt solicitation submitted by non-management Acc-no: 0001214659-21-011362 (34 Act) Size: 98 KB	2021-11-10	001-37845 211393647
DEFA14A	Documents	Additional definitive proxy soliciting materials and Rule 14(a)(12) material Acc-no: 0001193125-21-313501 (34 Act) Size: 16 KB	2021-10-29	001-37845 211364076

Filing Details (1)

- Since the early 2000s filings consist of multiple files.
- More and more filings are in html.

Filing Detail

SEC Home » Search the Next-Generation EDGAR System » Company Search » Current Page

Form 10-Q - Quarterly report [Sections 13 or 15(d)]

Filing Date 2015-04-23	Period of Report 2015-03-31
Accepted 2015-04-23 16:09:01	Filing Date Changed 2015-04-23
Documents 12	

[Interactive Data](#)

Document Format Files

Seq	Description	Document	Type	Size
1	FORM 10-Q	d860721d10q.htm	10-Q	1424179
2	EXHIBIT 15	d860721dex15.htm	EX-15	2622
3	EXHIBIT 31.1	d860721dex311.htm	EX-31.1	5767
4	EXHIBIT 31.2	d860721dex312.htm	EX-31.2	5761
5	EXHIBIT 32.1	d860721dex321.htm	EX-32.1	2941
6	EXHIBIT 32.2	d860721dex322.htm	EX-32.2	2930
	Complete submission text file	0001193125-15-144151.txt		14711709

Period of Report = fiscal quarter/year end


Complete submission file; always in txt format but with lots of html code!

Main document

Exhibits, e.g., certifications, pictures, press releases, etc.

Filing Details (2)

- Before 2000 only complete submission file.
- Hardly any html code.



Filing Detail

[SEC Home](#) » [Search the Next-Generation EDGAR System](#) » [Company Search](#) » [Current Page](#)

Form 10-Q - Quarterly report [Sections 13 or 15(d)]

Filing Date 1999-11-05	Period of Report 1999-09-30
Accepted 1995-09-28 00:00:00	
Documents 2	

Document Format Files

Seq	Description	Document
	Complete submission text file	0001032210-99-001526.txt

MICROSOFT CORP (Filer) CIK: 0000789019 (see all company filings)

IRS No.: 911144442 | State of Incorpor.: WA | Fiscal Year End: 0630
 Type: 10-Q | Act: 34 | File No.: 000-14278 | Film No.: 99741829
 SIC: 7372 Services-Prepackaged Software
 Assistant Director 3



Firm identifiers in EDGAR (1)

CIK (Central Index Key)

- 10-digit number, e.g., 0000789019 ↔ Microsoft Corp.
 - Unique firm ID
 - One CIK refers to exactly one firm.
 - However, one firm have multiple CIKs.
 - The CIK is available in Compustat
 1. Merge CRSP with Compustat
 2. Merge with SEC data
- **Problem: only today's CIK available in Compustat → BIAS!**

Problem: current CLKs (1)

General Motors Corp (permno 12079, gvkey 005073)

Filed for bankruptcy under chapter 11 on June 8, 2009. Reemerged on July 10, 2009

Data Editor (Browse) - [CRSP_monthly_1970_2014_with_gvkey.dta]

File Edit View Data Tools

ncusip[74228] 37045V10

	permno	date	gvkey	comnam	shrcd	exchcd	siccd	ncusip	ticker	permco	cusip	dlstcd	dlprc	dlret
67255	12079	31mar2008	005073	GENERAL MOTORS CORP	1.1e+01	1.0e+00	3711	37044210	GM	20799	37044210	.	.	.
67256	12079	30apr2008	005073	GENERAL MOTORS CORP	1.1e+01	1.0e+00	3711	37044210	GM	20799	37044210	.	.	.
67257	12079	30may2008	005073	GENERAL MOTORS CORP	1.1e+01	1.0e+00	3711	37044210	GM	20799	37044210	.	.	.
67258	12079	30jun2008	005073	GENERAL MOTORS CORP	1.1e+01	1.0e+00	3711	37044210	GM	20799	37044210	.	.	.
67259	12079	31jul2008	005073	GENERAL MOTORS CORP	1.1e+01	1.0e+00	3711	37044210	GM	20799	37044210	.	.	.
67260	12079	29aug2008	005073	GENERAL MOTORS CORP	1.1e+01	1.0e+00	3711	37044210	GM	20799	37044210	.	.	.
67261	12079	30sep2008	005073	GENERAL MOTORS CORP	1.1e+01	1.0e+00	3711	37044210	GM	20799	37044210	.	.	.
67262	12079	31oct2008	005073	GENERAL MOTORS CORP	1.1e+01	1.0e+00	3711	37044210	GM	20799	37044210	.	.	.
67263	12079	28nov2008	005073	GENERAL MOTORS CORP	1.1e+01	1.0e+00	3711	37044210	GM	20799	37044210	.	.	.
67264	12079	31dec2008	005073	GENERAL MOTORS CORP	1.1e+01	1.0e+00	3711	37044210	GM	20799	37044210	.	.	.
67265	12079	30jan2009	005073	GENERAL MOTORS CORP	1.1e+01	1.0e+00	3711	37044210	GM	20799	37044210	.	.	.
67266	12079	27feb2009	005073	GENERAL MOTORS CORP	1.1e+01	1.0e+00	3711	37044210	GM	20799	37044210	.	.	.
67267	12079	31mar2009	005073	GENERAL MOTORS CORP	1.1e+01	1.0e+00	3711	37044210	GM	20799	37044210	.	.	.
67268	12079	30apr2009	005073	GENERAL MOTORS CORP	1.1e+01	1.0e+00	3711	37044210	GM	20799	37044210	.	.	.
67269	12079	29may2009	005073	GENERAL MOTORS CORP	1.1e+01	1.0e+00	3711	37044210	GM	20799	37044210	.	.	.
67270	12079	30jun2009		GENERAL MOTORS CORP	1.1e+01	1.0e+00	3711	37044210	GM	20799	37044210	5.7e+02	.6100000143	-.186666653
74212	12369	30nov2010	005073	GENERAL MOTORS CO	1.1e+01	1.0e+00	3711	37045V10	GM	53554	37045V10	.	.	.
74213	12369	31dec2010	005073	GENERAL MOTORS CO	1.1e+01	1.0e+00	3711	37045V10	GM	53554	37045V10	.	.	.
74214	12369	31jan2011	005073	GENERAL MOTORS CO	1.1e+01	1.0e+00	3711	37045V10	GM	53554	37045V10	.	.	.
74215	12369	28feb2011	005073	GENERAL MOTORS CO	1.1e+01	1.0e+00	3711	37045V10	GM	53554	37045V10	.	.	.
74216	12369	31mar2011	005073	GENERAL MOTORS CO	1.1e+01	1.0e+00	3711	37045V10	GM	53554	37045V10	.	.	.
74217	12369	29apr2011	005073	GENERAL MOTORS CO	1.1e+01	1.0e+00	3711	37045V10	GM	53554	37045V10	.	.	.
74218	12369	31may2011	005073	GENERAL MOTORS CO	1.1e+01	1.0e+00	3711	37045V10	GM	53554	37045V10	.	.	.
74219	12369	30jun2011	005073	GENERAL MOTORS CO	1.1e+01	1.0e+00	3711	37045V10	GM	53554	37045V10	.	.	.
74220	12369	29jul2011	005073	GENERAL MOTORS CO	1.1e+01	1.0e+00	3711	37045V10	GM	53554	37045V10	.	.	.
74221	12369	31aug2011	005073	GENERAL MOTORS CO	1.1e+01	1.0e+00	3711	37045V10	GM	53554	37045V10	.	.	.
74222	12369	30sep2011	005073	GENERAL MOTORS CO	1.1e+01	1.0e+00	3711	37045V10	GM	53554	37045V10	.	.	.
74223	12369	31oct2011	005073	GENERAL MOTORS CO	1.1e+01	1.0e+00	3711	37045V10	GM	53554	37045V10	.	.	.
74224	12369	30nov2011	005073	GENERAL MOTORS CO	1.1e+01	1.0e+00	3711	37045V10	GM	53554	37045V10	.	.	.
74225	12369	30dec2011	005073	GENERAL MOTORS CO	1.1e+01	1.0e+00	3711	37045V10	GM	53554	37045V10	.	.	.



Problem: current CIKs (2)

General Motors Corp (permno 12079, gvkey 005073)

- GM was assigned a new CIK after bankruptcy.
- Not accounting for the CIK change will result in missing EDGAR data for GM for the entire period from 1997 – 2009!

	gvkey	datadate	cik	tic	fyear
67132	005073	31dec2003	0001467858	GM	2003
67133	005073	31dec2004	0001467858	GM	2004
67134	005073	31dec2005	0001467858	GM	2005
67135	005073	31dec2006	0001467858	GM	2006
67136	005073	31dec2007	0001467858	GM	2007
67137	005073	31dec2008	0001467858	GM	2008
67138	005073	31dec2009	0001467858	GM	2009
67139	005073	31dec2010	0001467858	GM	2010
67140	005073	31dec2011	0001467858	GM	2011
67141	005073	31dec2012	0001467858	GM	2012
67142	005073	31dec2013	0001467858	GM	2013

Compustat data for GM

EDGAR Search Results

SEC Home » Search the Next-Generation EDGAR System » Company Search » Current Page

General Motors Co CIK#: 0001467858 (see all company filings)

SIC: 3711 - MOTOR VEHICLES & PASSENGER CAR BODIES
 State location: MI | Fiscal Year End: 1231
 (Assistant Director Office: 5)
 Get **insider transactions** for this issuer.
 Get **insider transactions** for this reporting owner.

Filter Results: Filing Type: Prior to: (YYYYMMDD)

Filings	Format	Description
---------	--------	-------------

<http://www.sec.gov/cgi-bin/browse-edgar> **No search results!**

Home | Search the Next-Generation EDGAR System | Previous Page

Problem: current CIKs (3)

- General Motors Corp (permno 12079, gvkey 005073)



EDGAR Search Results

[SEC Home](#) » [Search the Next-Generation EDGAR System](#) » [Company Search](#) » [Current Page](#)

Companies with names matching "GENERAL MOTORS"

Click on CIK to view company filings

Items 1 - 11

CIK	Company
0000040729	Ally Financial Inc. SIC: 6172 - FINANCE LESSORS formerly: GENERAL MOTORS ACCEPTANCE CORP (filings through 2006-07-14) GMAC INC. (filings through 2010-05-07) GMAC LLC (filings through 2009-06-29)
0001037761	GENERAL MOTORS CAPITAL TRUST D
0001037734	GENERAL MOTORS CAPITAL TRUST G
0001467858	General Motors Co SIC: 3711 - MOTOR VEHICLES & PASSENGER CAR BODIES
0000040730	Motors Liquidation Co SIC: 3711 - MOTOR VEHICLES & PASSENGER CAR BODIES formerly: GENERAL MOTORS CORP (filings through 2009-07-13)
0001011448	GENERAL MOTORS EMPLOYES GLOBAL GROUP PENSION TRUST
0000804269	General Motors Financial Company, Inc. SIC: 6199 - FINANCE SERVICES formerly: AMERICREDIT CORP (filings through 2010-10-06)
0001497345	General Motors Holdings LLC
0000316750	GENERAL MOTORS HOURLY RATE EMPLOYEES PENSION TRUST ET AL
0000883422	GENERAL MOTORS INVESTMENT MANAGEMENT CORP formerly: Promark Investment Advisors (filings through 2010-08-03)
0000894123	GENERAL MOTORS INVESTMENT MANAGEMENT CORP GMIMCO /ADV

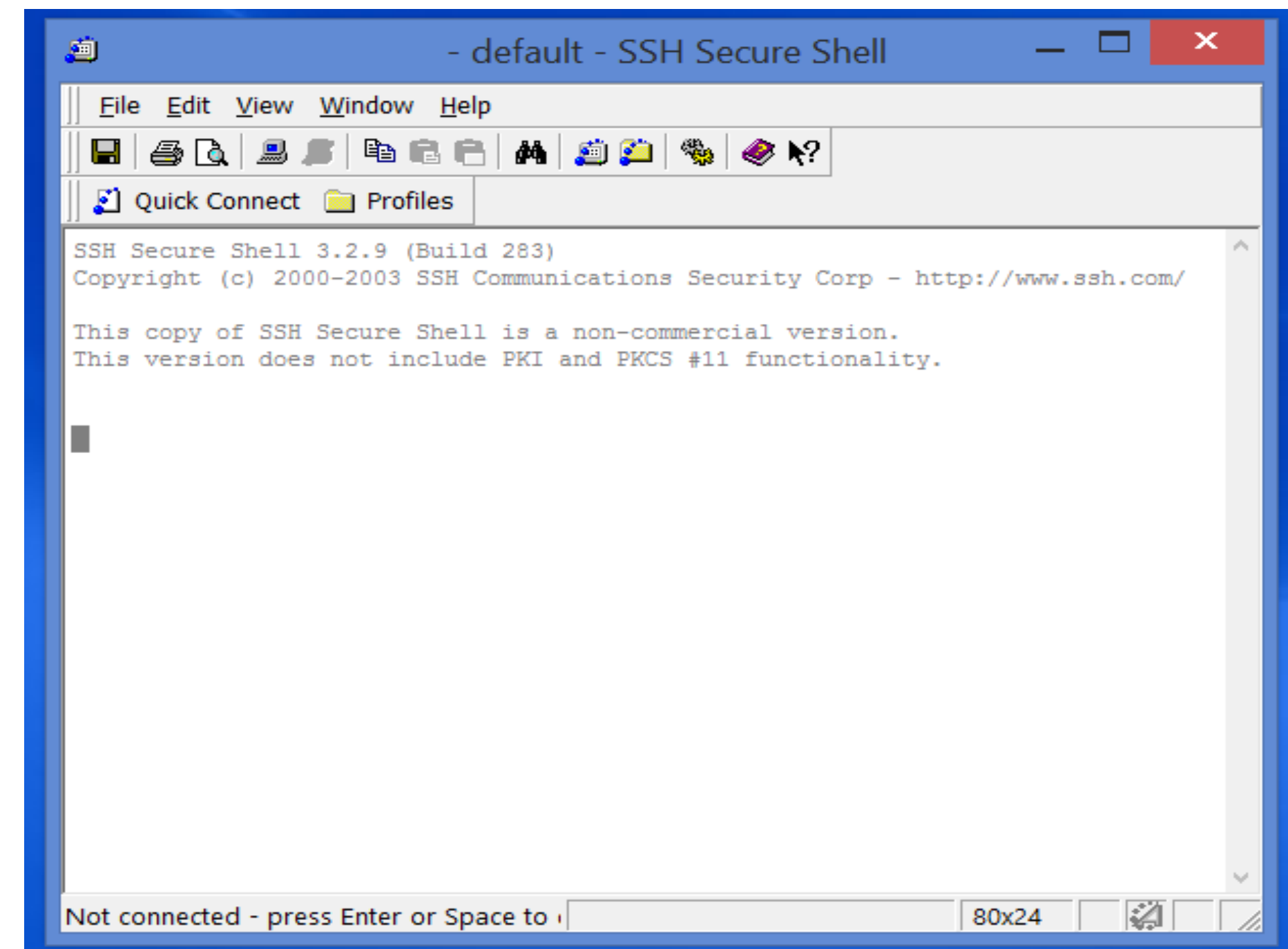
- Need to add old CIK
- “Motors Liquidation Co”
CIK==40730

Historical CIKs – Compustat Historical (1)

How to get historical CIKs?

Compustat Historical

- On the WRDS ftp server (not on the web interface)
- WRDS uses this data to create output files.
- Data only in SAS format.
- SSH Secure Shell needed to access the server



Historical CIKs – Compustat Historical (2)

Accessing Compustat Historical

```

File Edit View Window Help
[Icons]
Quick Connect Profiles New File Transfer Window

nedit <filename>          - Edit/create <filename> using nedit
pico <filename>          - Edit/create <filename> using pico
emacs <filename>         - Edit/create <filename> using emacs

qsub <filename>          - Submit a wrapper script for processing
qsas <filename>          - Submit a SAS job directly for processing
qrsh                      - Start an interactive session
qstat                    - Show your running and queued jobs
qstat -u \*              - Show all running and queued jobs
qghost -j                 - Show all grid hosts and thier jobs
qdel <job ID>            - Delete your job <job ID>

quota -s                 - Display available disk space
man <command>            - Read help/manual pages for <command>
menu                      - Show this menu
logout                    - Exit this session

For more information on the above, please reference the 'WRDS Cloud Manual' at:

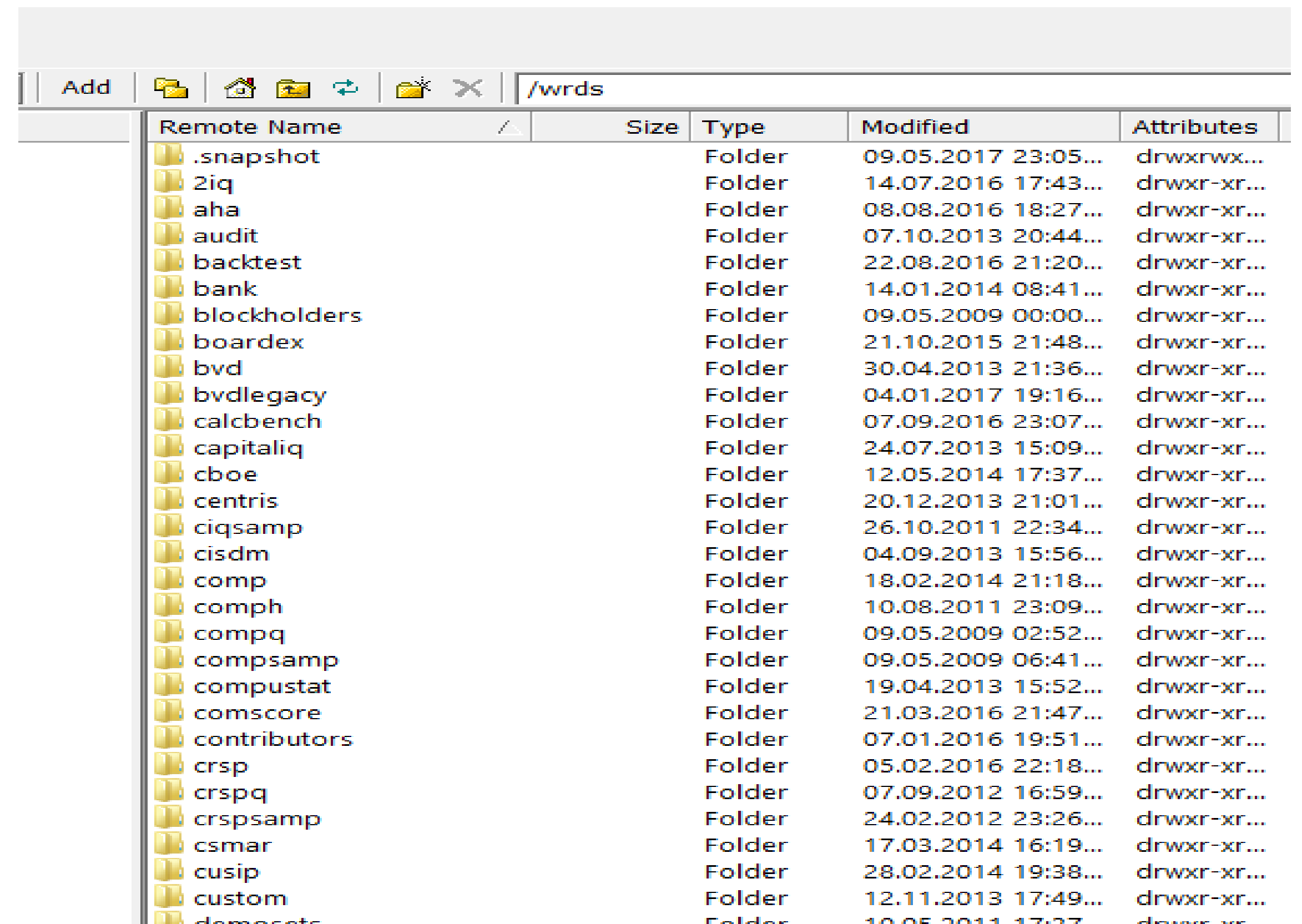
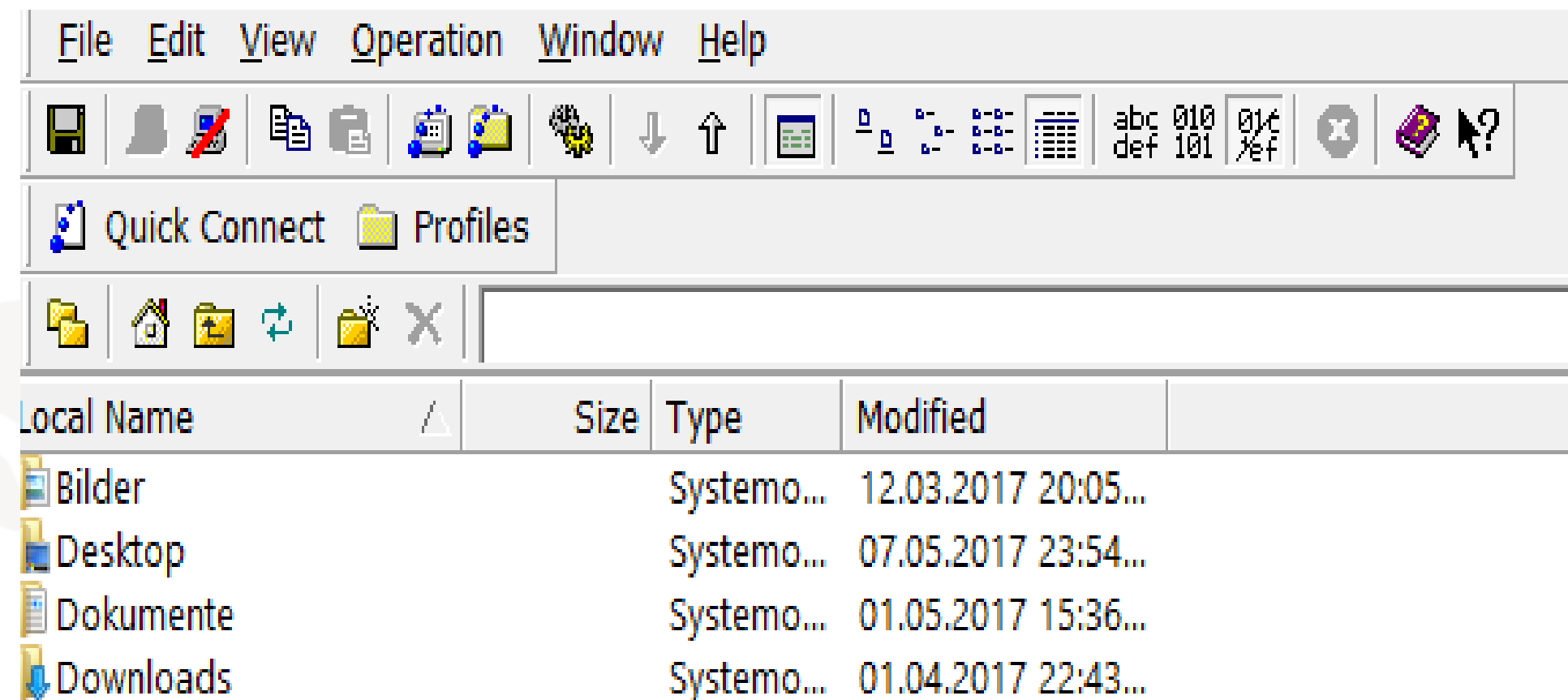
*      https://wrds-web.wharton.upenn.edu/wrds/support/      *

-----
[ahillert@wrds-cloud-login2-h ~]$
Open new file transfer window  SSH2 - aes128-cbc - hmac-md5  80x24

```

Historical CIKs – Compustat Historical (3)

Accessing Compustat Historical – File transfer window



Historical CIKs – Compustat Historical (4)

Accessing Compustat Historical – File transfer window - continued

The screenshot shows a file transfer window with a toolbar at the top containing icons for 'Add', home, refresh, and close. The address bar displays the path `/wrds/crsp/sasdata/a_ccm`. Below the toolbar is a table listing files with columns for Remote Name, Size, Type, Modified, and Attributes.

Remote Name	Size	Type	Modified	Attributes
ccm_lookup.sas7bdat	14,024,7...	SAS7BD...	01.02.2017 22:05...	-rw-r-----
ccm_lookup.sas7bndx	7,118,848	SAS7BN...	01.02.2017 22:05...	-rw-r-----
ccm_qvars.sas7bdat	196,608	SAS7BD...	01.02.2017 22:05...	-rw-r-----
ccmxfpf_linktable.sas7bd...	4,325,376	SAS7BD...	01.02.2017 22:05...	-rw-r-----
ccmxfpf_linktable.sas7bn...	4,743,168	SAS7BN...	01.02.2017 22:05...	-rw-r-----
ccmxfpf_lnkhist.sas7bdat	4,915,200	SAS7BD...	01.02.2017 22:05...	-rw-r-----
ccmxfpf_lnkhist.sas7bndx	7,462,912	SAS7BN...	01.02.2017 22:05...	-rw-r-----
ccmxfpf_lnkrng.sas7bdat	26,476,5...	SAS7BD...	01.02.2017 22:05...	-rw-r-----
ccmxfpf_lnkrng.sas7bndx	4,014,080	SAS7BN...	01.02.2017 22:05...	-rw-r-----
ccmxfpf_lnkused.sas7bdat	5,898,240	SAS7BD...	01.02.2017 22:05...	-rw-r-----
ccmxfpf_lnkused.sas7bndx	7,471,104	SAS7BN...	01.02.2017 22:05...	-rw-r-----
comphead.sas7bdat	142,868,...	SAS7BD...	01.02.2017 22:05...	-rw-r-----
comphead.sas7bndx	2,555,904	SAS7BN...	01.02.2017 22:05...	-rw-r-----
comphist.sas7bdat	945,291,...	SAS7BD...	01.02.2017 22:05...	-rw-r-----
comphist.sas7bndx	16,392,1...	SAS7BN...	01.02.2017 22:05...	-rw-r-----
compmaster.sas7bdat	3,670,016	SAS7BD...	01.02.2017 22:05...	-rw-r-----
cst_hist.sas7bdat	74,383,3...	SAS7BD...	01.02.2017 22:05...	-rw-r-----
cst_hist.sas7bndx	3,424,256	SAS7BN...	01.02.2017 22:05...	-rw-r-----
sechead.sas7bdat	7,471,104	SAS7BD...	01.02.2017 22:05...	-rw-r-----
sechead.sas7bndx	5,177,344	SAS7BN...	01.02.2017 22:05...	-rw-r-----
sechist.sas7bdat	15,728,6...	SAS7BD...	01.02.2017 22:05...	-rw-r-----
sechist.sas7bndx	11,427,8...	SAS7BN...	01.02.2017 22:05...	-rw-r-----

Historical CIKs – Compustat Historical (5)

Compustat Historical

- General Motors (CIKs: 0000040730 and 0001467858)

hchgdt	hchgnddt	hcik	hconm	hconml	gvkey
14apr2007	11apr2008	0000040730	GENERAL MOTORS CORP	General Motors Corp	005073
12apr2008	12jun2009	0000040730	GENERAL MOTORS CORP	General Motors Corp	005073
13jun2009	31jul2009	0000040730	GENERAL MOTORS CORP	General Motors Corp	005073
01aug2009	07may2010	0000040730	MOTORS LIQUIDATION CO	Motors Liquidation Co	005073
08may2010	10dec2010	0001467858	MOTORS LIQUIDATION CO	Motors Liquidation Co	005073
11dec2010	13may2011	0001467858	GENERAL MOTORS CO	General Motors Co	005073
14may2011	09mar2012	0001467858	GENERAL MOTORS CO	General Motors Co	005073
10mar2012	13apr2012	0001467858	GENERAL MOTORS CO	General Motors Co	005073
14apr2012	08mar2013	0001467858	GENERAL MOTORS CO	General Motors Co	005073
09mar2013	07mar2014	0001467858	GENERAL MOTORS CO	General Motors Co	005073
08mar2014	13feb2015	0001467858	GENERAL MOTORS CO	General Motors Co	005073
14feb2015	13mar2015	0001467858	GENERAL MOTORS CO	General Motors Co	005073
14mar2015		0001467858	GENERAL MOTORS CO	General Motors Co	005073

--> CIK Wechsel am 08.05.2010

weitgehend zu CRSP: Bankruptcy in Jun 2009 und Re-emergence in Nov 2010

permno	date	ncusip	comnam	permco	dlstcd	dlprc	dlret
12079	30apr2009	37044210	GENERAL MOTORS CORP	20799			
12079	29may2009	37044210	GENERAL MOTORS CORP	20799			
12079	30jun2009	37044210	GENERAL MOTORS CORP	20799	5.7e+02	.6100000143	-.186666653
...							
permno	date	ncusip	comnam	permco	prc	vol	ret
12369	30nov2010	37045V10	GENERAL MOTORS CO	53554	34.20000076	7575358	
12369	31dec2010	37045V10	GENERAL MOTORS CO	53554	36.86000061	3726099	.077777773
12369	31jan2011	37045V10	GENERAL MOTORS CO	53554	36.49000168	3535732	-.010037952

File and folder structure in EDGAR (1)

SEC overview files on filed documents

- Go to 'full-index' folder: <https://www.sec.gov/Archives/edgar/full-index/>
- Access the data by year and quarter.

U.S. SECURITIES AND EXCHANGE COMMISSION

Search Options

ABOUT DIVISIONS ENFORCEMENT REGULATION EDUCATION FILINGS NEWS

NAME	SIZE	LAST MODIFIED
1993		05/06/2017 01:01:34 AM
1994		05/06/2017 01:03:33 AM
1995		05/06/2017 01:05:54 AM
1996		05/06/2017 01:08:34 AM
1997		05/06/2017 01:11:25 AM
1998		05/06/2017 01:14:18 AM
1999		05/06/2017 01:17:18 AM
2000		05/06/2017 01:20:26 AM

Parent Directory

NAME	SIZE	LAST MODIFIED
QTR1		09/06/2014 01:11:19 AM
QTR2		09/06/2014 01:11:48 AM
QTR3		09/06/2014 01:12:15 AM
QTR4		09/06/2014 01:12:48 AM

File and folder structure in EDGAR (2)

SEC overview file - form.idx

- The 'form.idx' provides you with a list of all filings submitted to the SEC in that quarter.
- Form.idx contains information on
 1. Filing Type
 2. Company Name
 3. CIK
 4. Filing Date
 5. Link to download the file
 - Add <http://www.sec.gov/Archives/> to the link!
- Form.idx are sorted by Filing Type (then by Company Name and Filing Date).

File and folder structure in EDGAR (3)

SEC overview file - How to get a comprehensive list of all submission?

- Download and append quarterly lists.
- This gives you an overview on:
 - Frequency of a specific filing type
 - Availability
 - Seasonality
- How to work with these lists?
 - The columns in the lists have a fixed length (# of characters).
 - E.g., import in Stata by “import delimited”.
 - Use substring function, e.g., generate filing_type=substr(v1,1,12) and generate company_name=substr(v1,13,60), etc.

Programming problem 2 asks you to search for and automatically download SEC filings.

File and folder structure in EDGAR (4)

SEC overview file - form.idx - daily

- In addition to the quarterly files, there are daily files.
- <https://www.sec.gov/Archives/edgar/daily-index/>
- Problems of daily files
 - Inconsistent file names, e.g., YYYYMMDD vs. YYMMDD vs. MMDDYY.
 - In Q2 2014, there are only zipped files available.
 - Duplicate entries (even within the same file).
Duplicates = same CIK + filing type + filing date + link to EDGAR (incl. filename).

NAME	SIZE	LAST MODIFIED
company.20001002.idx	120 KB	01/24/2003 02:23:22 PM
company.20001003.idx	121 KB	01/24/2003 02:23:40 PM
company.20001004.idx	108 KB	01/24/2003 02:24:02 PM
company.20001005.idx	109 KB	01/24/2003 02:24:24 PM
company.20001006.idx	147 KB	01/24/2003 02:24:50 PM
company.20001010.idx	311 KB	01/24/2003 02:25:23 PM
company.20001011.idx	126 KB	01/24/2003 02:25:51 PM
company.20001012.idx	109 KB	01/24/2003 02:26:24 PM
company.20001013.idx	145 KB	01/24/2003 02:26:59 PM
company.20001016.idx	135 KB	01/24/2003 02:27:24 PM

File and folder structure in EDGAR (5)

Comparison of quarterly and daily overview files

- Approach: merge list of all filings based on filing type, filing date, CIK, and link on EDGAR.
- Sample period: 1994Q1 to 2020Q1
- Merged: 19,354,372
- Not matched:
 - from quarterly 217,379
 - Most years have 1,000 to 4,000 unmatched observations.
 - 1994 (50,192), 1996 (37,695), 1999 (29,619), 1997 (21,444), and 1995 (20,684) account for over 70% of non-matches.
 - Almost full match in 1998 (93).
 - from daily 351,075
 - Most years have 3,000 to 6,000 unmatched observations.
 - 1996 (57,940), 1998 (56,969), 1997 (37,567), 1999 (29,811), 1994 (28,678), and 1995 (27,504) account for about 67% of non-matches.

File and folder structure in EDGAR (6)

Comparison of quarterly and daily overview files - continued

- Looking at filing type does not show any pattern.
- Looking at 10-Ks (10-K, 10-K405, 10KSB) results in
 - Merged: 249,665
 - Unmatched observations from quarterly 3,603
 - Most years have 20 to 30 unmatched observations.
 - Early years + 2005 to 2007 (147 to 254 non-matches) account for most observations.
 - Almost full match ($N < 4$) in 1998, 2003, and 2004.
 - Unmatched observations from daily 4,690
 - Most years have 20 to 30 unmatched observations.
 - Period up to 2007 (except for 03 and 04) show several hundred non-matches.
- Recommendation: take combination of both lists.

File and folder structure in EDGAR (7)

Using the link in the form.idx

- The link in the form.idx file relates to the complete submission file.
- Example
 - 10-K of AMAZON COM INC (CIK 1018724) filed on 2008-02-11.
 - Link: edgar/data/1018724/0001193125-08-024707.txt
 - complete link: <http://www.sec.gov/Archives/edgar/data/1018724/0001193125-08-024707.txt>
- To get to the overview of the submitted files.
 - Replace the '.txt' by '-index.htm'
 - <http://www.sec.gov/Archives/edgar/data/1018724/0001193125-08-024707-index.htm>
 - You can access the individual files of the submission directly. For example,
 - Main document's name: 'd10k.htm'
 - Delete all '-' from the original link.
 - Replace the '.txt' by '/d10k.htm'
 - <http://www.sec.gov/Archives/edgar/data/1018724/000119312508024707/d10k.htm>

Data availability (1)

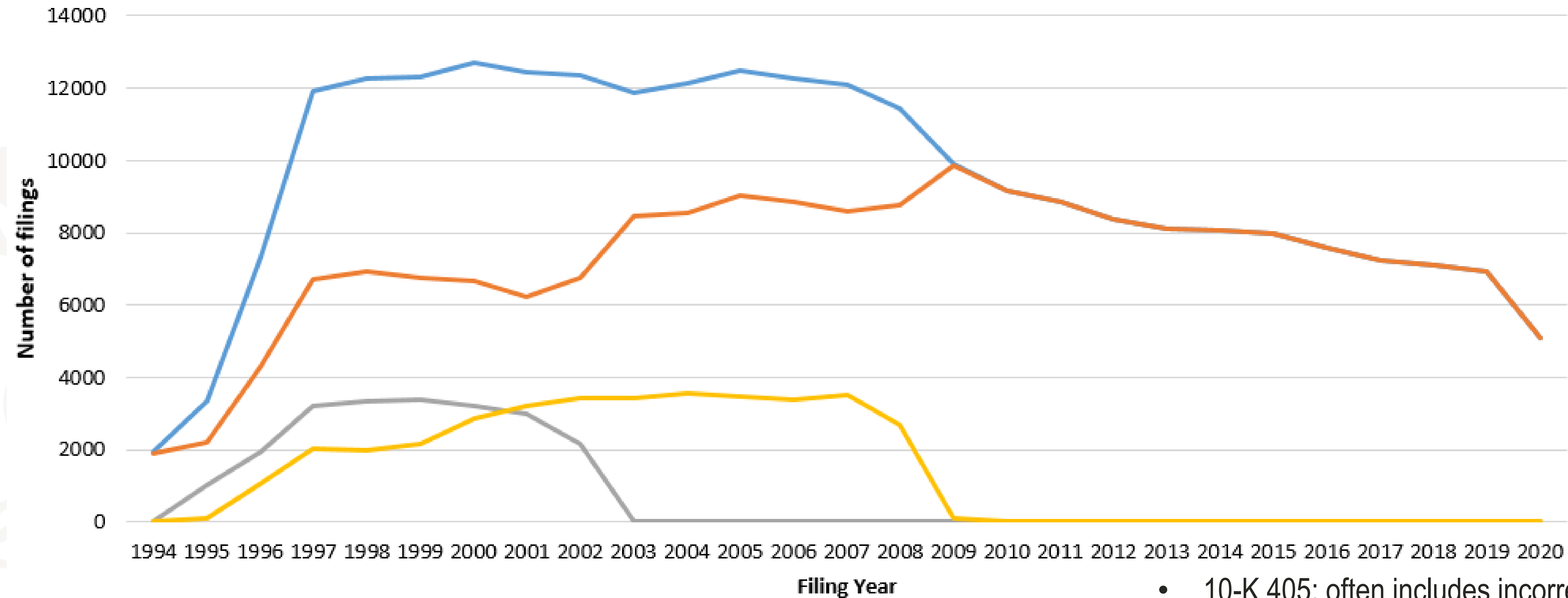
Data availability in EDGAR

- EDGAR starts in 1994
- However, firms were not required to submit digital forms.
 - *'We begin in 1997 because this is the first year all firms are required to file digital forms.'*
Loughran and McDonald (2011, footnote 18)
 - *'Since electronic filings on EDGAR were not required until 1996, the early sample is tilted towards larger, more sophisticated companies. Larger firms tend to have longer 10-Ks.'*
Loughran and McDonald (2015, note 5)
 - *'Some large firms begin electronic filing in 1994. Electronic filing was required of all firms, with minor exceptions, beginning in 1996.'*
Loughran and McDonald (2016, footnote 2).

Data availability (2)

Number of 10-Ks over time

Number of 10-K related filings over time

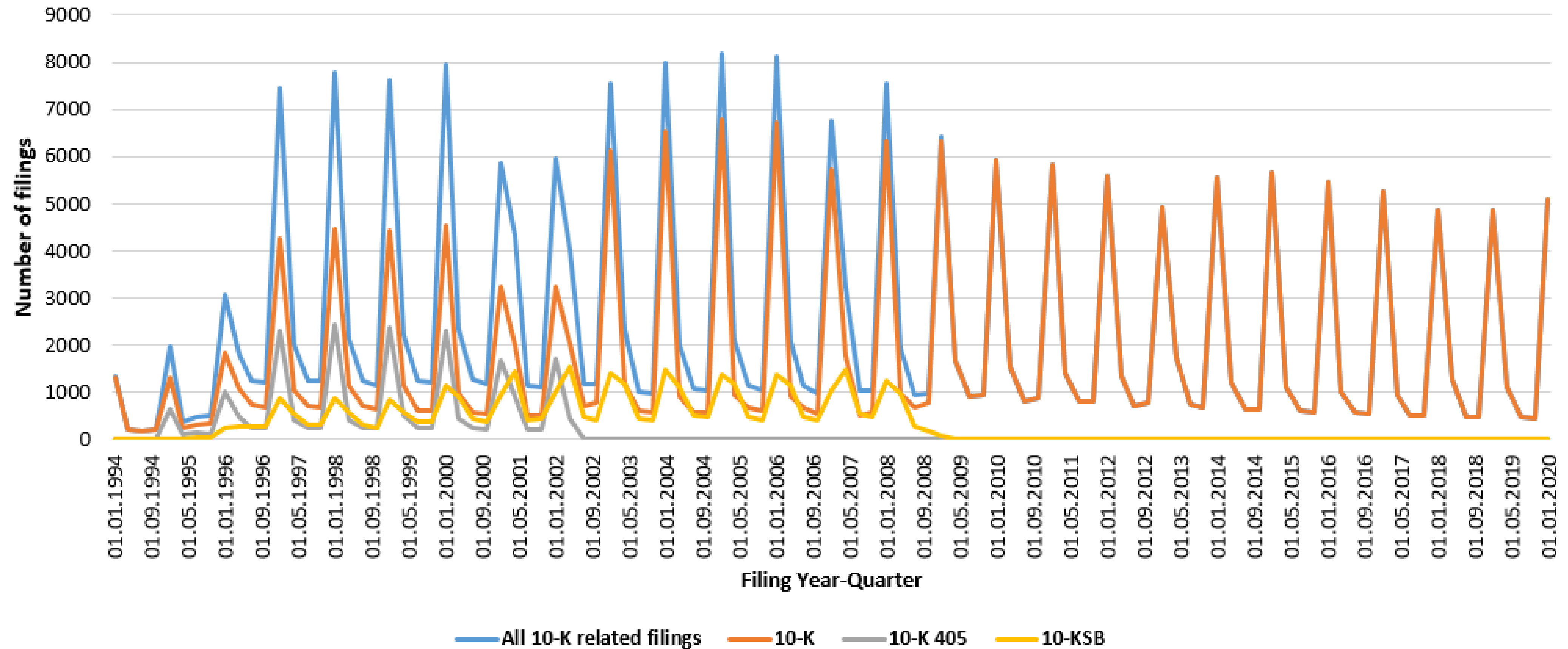


- 10-K 405: often includes incorrectly classified 10-Ks (see footnote 4 in Loughran and McDonald, 2011).
- 10-KSB: small businesses

Data availability (3)

Seasonality

Number of 10-K related filings per year-quarter



Firm identifiers in EDGAR (2)

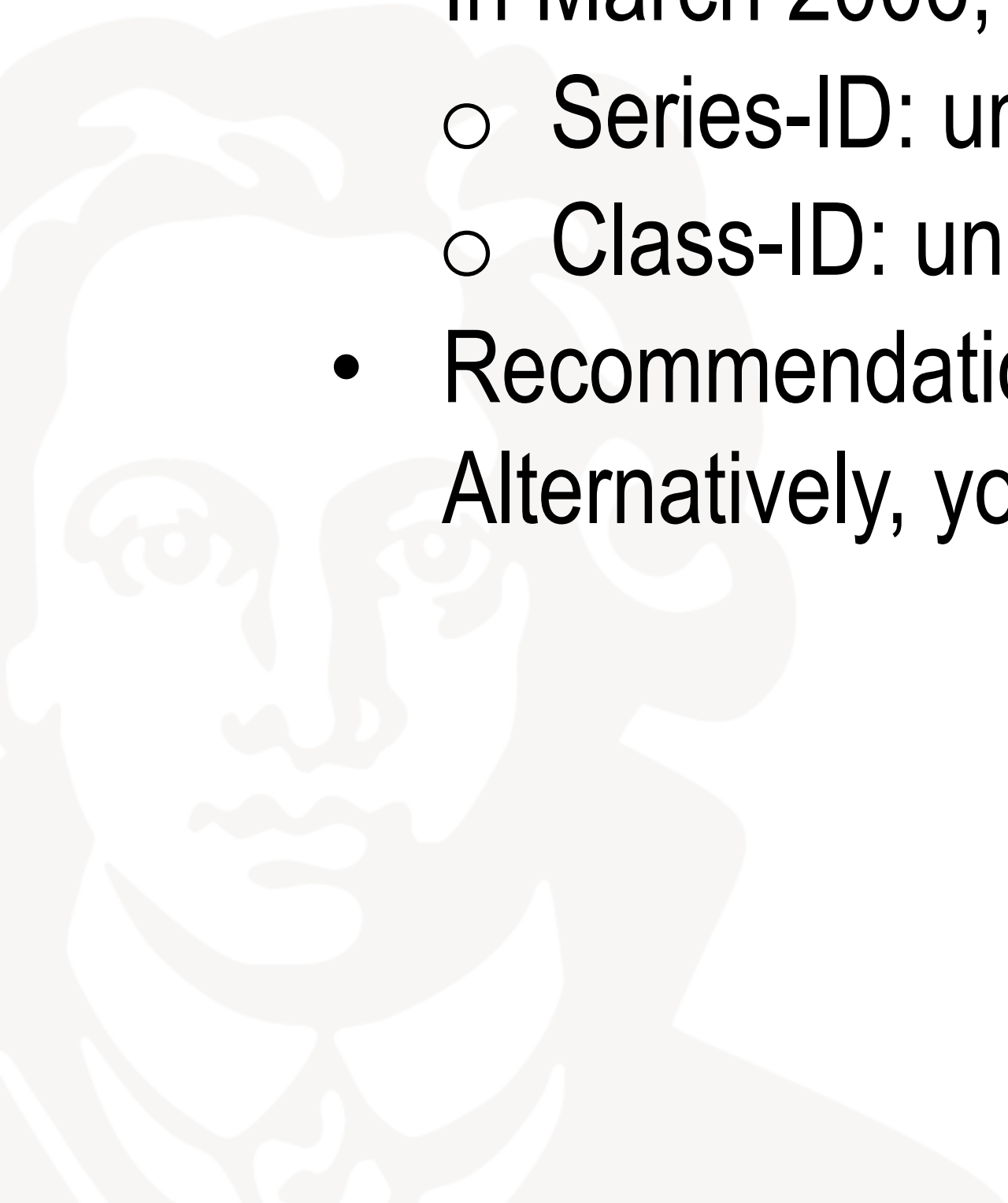
Ticker in EDGAR

- Search shows the filings for the stock that currently trades under the ticker symbol.
 - Example: 'AAC'
 - AAC Holdings Inc (permno 14945, Oct 2014 - Dec 2014)
 - Arcadia Financial LTD (permno 77297, May 1997 – Apr 2000)
 - AbleAuctions.Com Inc (permno 88278, Jun 2000 – Jan 2010)
- When using firms' ticker symbols, pay attention to the corresponding time period.

Firm identifiers in EDGAR (3)

Identifiers for funds and investment companies

- CLK is a unique investment management company identifier.
- CLK can (but need not!) refer to a single fund, e.g., 0000061397 → Fidelity Magellan Fund
- In March 2006, the SEC introduced additional IDs:
 - Series-ID: unique portfolio (=fund) identification number.
 - Class-ID: unique share class identification number.
- Recommendation: merge with CRSP mutual fund and MFLinks by ticker.
Alternatively, you can use name matching.



3. Other databases



Further data sources

Transcripts of earnings conference calls

- There are several sources
 - SeekingAlpha.com collects conference call transcripts.
http://seekingalpha.com/earnings_center/transcripts
You have to register (for free) to get full access.
 - LexisNexis FD Newswire
 - Thomson Reuters/Refinitiv
- Rapidly growing research interest in earnings calls. Prior research shows that the qualitative information is informative for future firm performance.

Analyst reports

- Recommendations/estimates (without reports): I/B/E/S
- Full text available via Thomson One Banker.
- Tone of analyst reports increases the explanatory power for abnormal stock returns around the publication date.

What have we learnt in this part?

What have we learnt in this part?

- LexisNexis
 - Comprehensive coverage
 - Relevance score
- SEC / EDGAR
 - Correctly linking CRSP/Compustat to EDGAR is difficult.
 - Historical CIKs are needed. Otherwise, systematic bias!
- Verbal communication: earnings conference calls

Thank you very much for your attention!

