

The background of the slide features a large, faint, circular seal of Rutgers University. The seal contains the text 'RUTGERS UNIVERSITY' and 'THE STATE UNIVERSITY OF NEW JERSEY' around its perimeter, with a central sunburst design.

RUTGERS

THE STATE UNIVERSITY
OF NEW JERSEY

Continuous Auditing & Reporting Laboratory (CARLab) Research Overview

As of June 2023

<u>Tracking Technologies: An Examination of Hospital Privacy Compliance and IT System Control Effectiveness</u>	<u>Whether Firms' Risk Disclosures are Still Informative? Using a Novel Tool to Measure Textual Data</u>	<u>Can External Auditors Utilize ChatGPT to Assure ESG Reports? - Starting from Environmental Section Test</u>	<u>CONTINUOUS MONITORING SYSTEM WITH INTERACTIVE VISUALIZATION: CASE STUDY ON GOVERNMENT PAYROLL DATABASE</u>	<u>The Relationship Between the Proportion of Outliers and the Average Attenuation Rate of Outlier Scores: An Approach of Algorithm Selection in Unsupervised Outlier Detection without Ground Truth</u>
<u>The benefits of big data: the case of wind data to detect greenwashing</u>	<u>The Watchdog Effect of Granular Transparency Regulation: Evidence from Firms' Responses to the Greenhouse Gas Reporting Program</u>	<u>Sentimental Analysis as a measure for determining the Informativeness of Cybersecurity Disclosures</u>	<u>Enhancing Payroll Fraud Detection for Governments: Continuous Auditing with Deep Learning based Outlier Detection</u>	<u>RPG gamification in accounting education</u>
<u>Partner-client Social Ties and Audit Quality.</u>	<u>Qualitative accounting education research using ChatGPT.</u>	<u>Analyzing Analyst Question Patterns in Earnings Calls: Leveraging NLP Models for Sentiment and Forward-Looking Analysis</u>	<u>ESG violation and insider trading</u>	<u>The Usefulness of Large Language Models in External Audits</u>
<u>Governance and Donations: Evidence from Small Nonprofit Organizations</u>	<u>AI Co-pilot auditing Plus More</u>	<u>Do employees know what would happen? Signals for the firm performance derived from the bottom tone</u>	<u>Municipality Government Reporting Automation</u>	<u>Bad Debt In Emergency Department</u>
<u>ESG - Deloitte</u>	<u>Manual Journal Entry Risk Identification Using Natural Language Processing and Deep Learning</u>	<u>ESG Report Data Extraction: A Generalized Text-Mining Framework.</u>	<u>Predicting revenues using customer reviews</u>	
<u>The Increased Role of Advanced Technology and Automation in Audit A</u>	<u>Collaborate Verbal, Voice, and Visual Evidence in Detecting Managerial Uncertainty and Financial</u>	<u>AIS Literature Review Generation</u>	<u>Continuous Auditing in the Age of AI: An Exploration of OpenAI's GPT Model</u>	

The background of the slide features a large, faint watermark of the Rutgers University seal. The seal is circular with a sunburst in the center and the words "RUTGERS UNIVERSITY" around the perimeter.

RUTGERS

THE STATE UNIVERSITY
OF NEW JERSEY

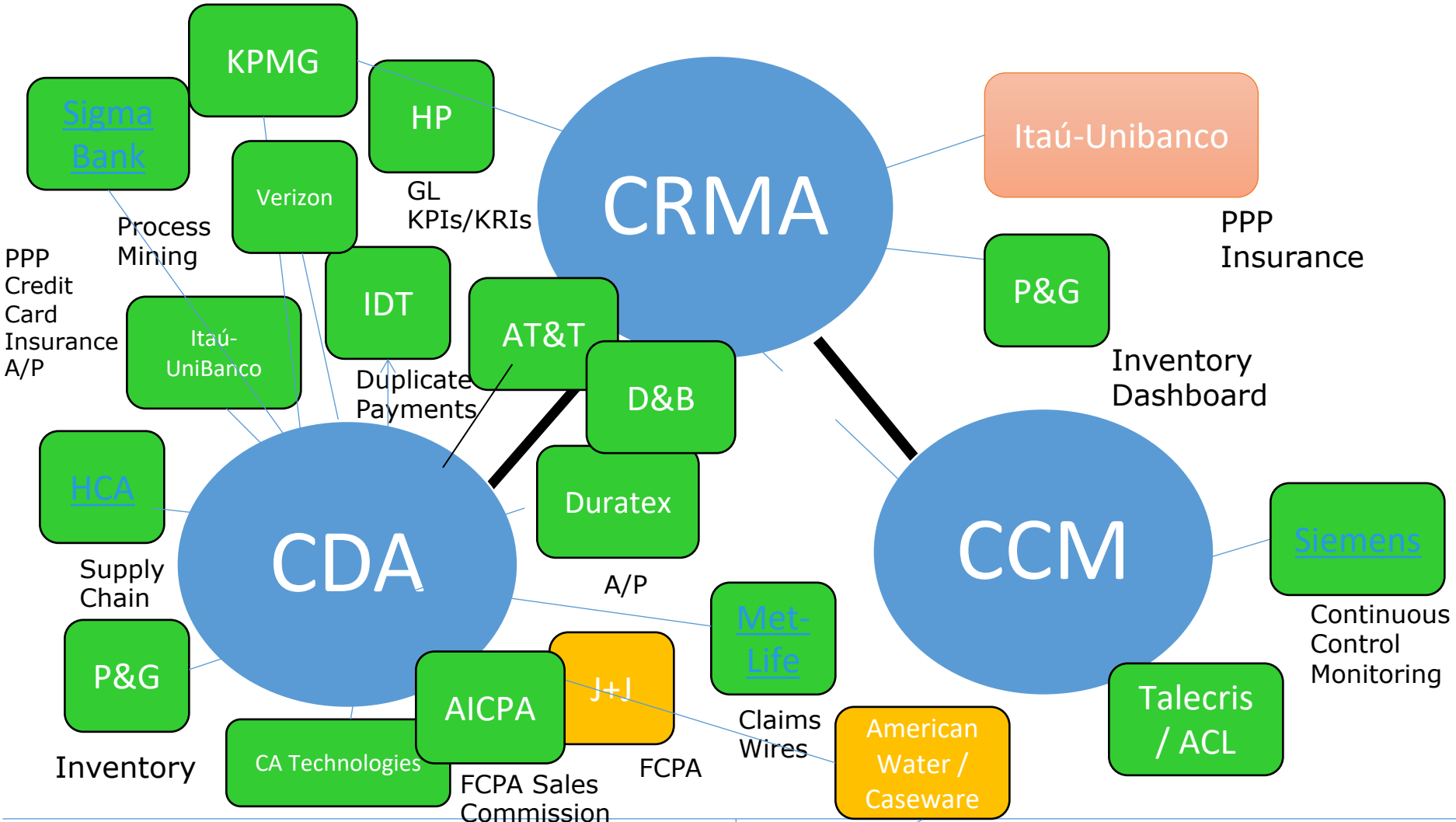
The CarLab

Continuous Audit and Reporting Laboratory

- Graduate School of Management
- Rutgers University



CarLab 2017



Audit Automation

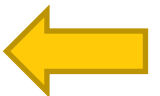
- P&G: Order to Cash**
- Auditor Judgment**
- Siemens- AAS Automation**
- AICPA – ADS / APS**

Audit Methodologies

- **Multidimensional Clustering**
- **Process Mining**
- **Continuity Equations**
- **Predictive Auditing**
- **Visualization**
- **Analytic Playpen**

Audit Innovation at the CarLab

- Continuous Audit at AT&T
- Continuity Equations
- Multidimensional clustering
- Process mining
- Predictive auditing
- Conceptualization of the distance audit
- Rule based outlier detection
- Exploratory data analysis
- Exceptional Exceptions
- Audit Data Standards
- Audit Apps
- Visualization
- Deep Learning and Machine Learning in audit
- Cognitive computing and aid for audit brainstorm
- Blockchain and Smart Contract
- Robotic Process Automation
- Intelligent Process Automation



The background of the slide features a large, faint watermark of the Rutgers University Newark seal. The seal is circular with a sunburst design in the center and the words "RUTGERS UNIVERSITY" around the perimeter. In the top left corner, the word "RUTGERS" is written in a large, red, serif font.

RUTGERS

RUTGERS
UNIVERSITY | NEWARK

Tracking Technologies: An Examination of Hospital Privacy Compliance and IT System Control Effectiveness

Hilal Atasoy

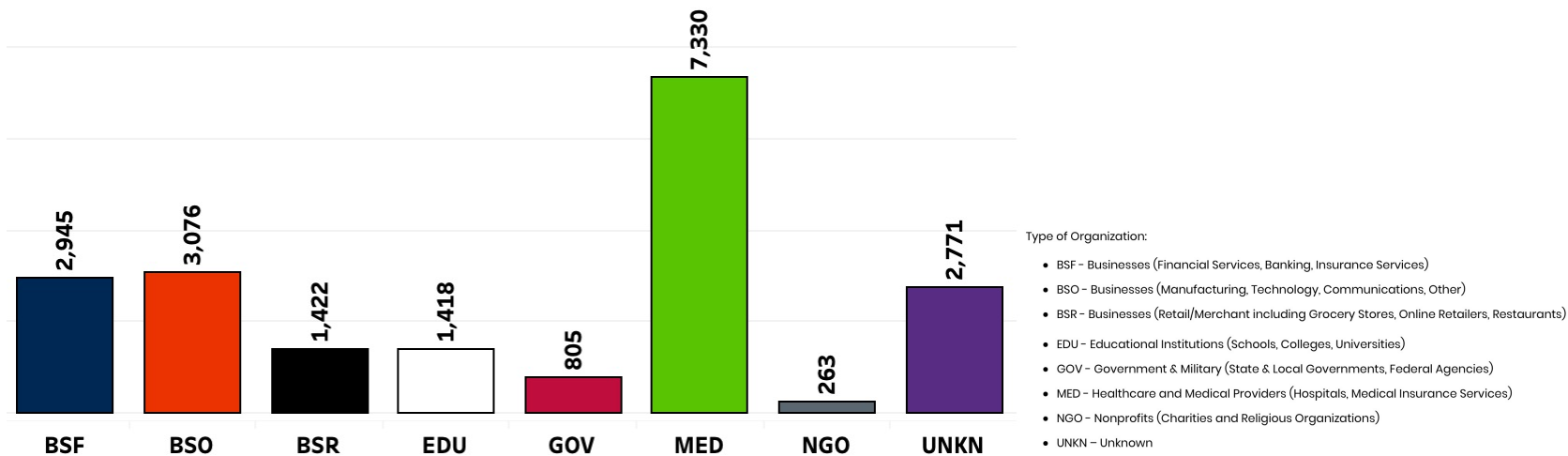
Ryan McDonough

Guangyue (Maria) Zhang

Rutgers Business School

BACKGROUND – PRIVACY BREACH

- Privacyrights clearinghouse:



BACKGROUND - CONSUMER PRIVACY

- There is growing tension between ethical issues and a firm's economic interest raised by big data analytics and behavioral analysis that violates public norms for acceptable data use
- Consumer privacy protection regulations are essential to the change in privacy management practice
- EU General Data Protection Regulation (GDPR)
 - Proposed - 2012, agreed upon - December 2016, in effect - May 25, 2018
 - The scope, timing, means of enforcement, and business obligations
- What's new
 - On 04th January 2023, Ireland's Data Protection Commission fined Meta 390 million euros for breaches of the GDPR relating to its social media service
 - The platforms' basis for seeking user permission to collect data for personalized advertising is invalid

BUSINESS OBLIGATIONS	Notice/transparency requirements
	Legal basis for processing
	Purpose limitation
	Data minimization
	Security requirements
	Privacy by design
	Processor/service provider requirements
	Prohibition on discrimination
	Record keeping
	Risk/impact assessments
	Data breach notification*
	Registration with authorities
	Data protection officer
	International data transfer restrictions

BACKGROUND – TRACKING TECHNOLOGY

- It is a script or code on a website or mobile app used to gather information about users as they interact with the website or mobile app
- User information is then analyzed by owners of the website or mobile app or third parties to profile users (e.g., IP address) and create insights about users' online activities
 - Pro: improve user experience
 - Con: user profile and continuous tracking by third parties without patients' knowledge or consent
- The Meta Pixel collects sensitive health information and shares it with Facebook

```
document, 'script', 'https://web.archive.org/web/20170406180820/https://connect.facebook.net/en_US/fbevents.js');  
fbq('init', '1171011999584753');  
fbq('track', "PageView");  
</script>  
<noscript>
```

BACKGROUND - CONSUMER PRIVACY

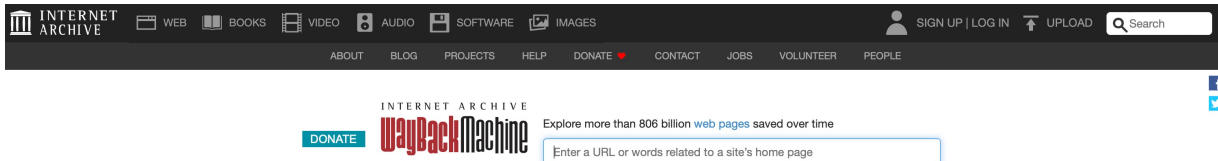
- U.S. federal consumer privacy law - Health Insurance Portability and Accountability Act (HIPAA) of 1996
 - **Objective:** Designed to protect patients' personal health information
 - **Requirements:** Obtaining authorization before use or disclosing, implement safeguards to protect the security health information
 - **Components:** Privacy rule, security rule, breach notification rule, and enforcement rule
 - **Subject:** entities(Healthcare providers, health plans, and healthcare clearinghouses) and business associates
- U.S. Department of Health & Human Services
 - The HIPAA Rules apply when the information that regulated entities collect through tracking technologies or disclose to tracking technology vendors includes protected health information (PHI).
 - Regulated entities are not permitted to use tracking technologies in a manner that would result in impermissible disclosures of PHI to tracking technology vendors or any other violations of the HIPAA Rules.

RESEARCH QUESTION

- Q: Does impermissible data sharing increase information security vulnerability?
- How do the HIPAA Rules apply to regulated entities' use of tracking technologies?
 - Tracking on user-authenticated webpages
 - Tracking on unauthenticated webpages
 - Tracking within mobile apps
- HIPAA compliance obligations for regulated entities when using tracking technologies
 - Regulated entities are required to comply with the HIPAA Rules when using tracking technologies.
- Online tracking technology - poor information security awareness and possible poor practice by the hospital management
 - The markup article of 2022 “Facebook Is Receiving Sensitive Medical Information from Hospital Websites” shows many hospital management are not aware of the code insertion.
 - Even if there is a vetting process of the code, there is no disclosure on the rigor of the process

DATA COLLECTION

- Hospital pixel tracking history
 - US News hospital list and website URLs
 - Internet Archive Wayback Machine: retrieve historical website HTML and java script



```
for y in range(2010, 2024):
    user_agent = "Mozilla/5.0 (Windows NT 5.1; rv:40.0) Gecko/20100101 Firefox/40.0"
    cdx = WaybackMachineCDXServerAPI(url, user_agent, start_timestamp=y, end_timestamp=y)
    for item in cdx.snapshots():
        url_list.append(item.archive_url)
    break
```

- Inspect retrieved code and script for “fbq()” scripts
- Final: 2010-2023 Hospital-year observation of pixel tracking usage

```
for i in tqdm(url_list):
    asession = AsyncHTMLSession()
    retry = Retry(connect=3, backoff_factor=0.5)
    adapter = HTTPAdapter(max_retries=retry)
    asession.mount('https://', adapter)
    r = await asession.get(i)
    await r.html.arender(timeout=10000)
    if 'fbq(' in str(r.html.raw_html):
        pixel_list.append(i)
        n+=1
    time.sleep(1)
    if (n % 2) == 0:
        time.sleep(1)
```

DATA COLLECTION

- Privacy breach
 - Privacyrights clearinghouse: breach chronology dataset, selecting “MED” organizations
 - Company name, Year of breach, Type of breach, Total Records, Date Made Public and description
 - Fuzzy matching of company name to merge the privacy breach dataset with the pixel tracking usage dataset
 - Final: 2010-2023 Hospital-year observation of pixel tracking usage indicator and breach incident indicator

	Hospital	Pixtrack_Year_Start	Score	year	pixel	Type of breach	Total Records	breach	pixel_byHospital	breach_byHospital
0	Mayo Clinic - Rochester	0.0	0.9796	2014	0	0	0	0.0	0.0	0.0
1	Mayo Clinic - Rochester	0.0	0.9796	2015	0	0	0	0.0	0.0	0.0
2	Mayo Clinic - Rochester	0.0	0.9796	2016	0	0	0	0.0	0.0	0.0
3	Mayo Clinic - Rochester	0.0	0.9796	2017	0	0	0	0.0	0.0	0.0
4	Mayo Clinic - Rochester	0.0	0.9796	2018	0	0	0	0.0	0.0	0.0
5	Mayo Clinic - Rochester	0.0	0.9796	2019	0	0	0	0.0	0.0	0.0
6	Mayo Clinic - Rochester	0.0	0.9796	2020	0	0	0	0.0	0.0	0.0
7	Cleveland Clinic	0.0	0.9439	2014	0	0	0	0.0	0.0	0.0
8	Cleveland Clinic	0.0	0.9439	2015	0	0	0	0.0	0.0	0.0
9	Cleveland Clinic	0.0	0.9439	2016	0	0	0	0.0	0.0	0.0
10	Cleveland Clinic	0.0	0.9439	2017	0	0	0	0.0	0.0	0.0
11	Cleveland Clinic	0.0	0.9439	2018	0	0	0	0.0	0.0	0.0
12	Cleveland Clinic	0.0	0.9439	2019	0	0	0	0.0	0.0	0.0
13	Cleveland Clinic	0.0	0.9439	2020	0	0	0	0.0	0.0	0.0
14	Massachusetts General Hospital	2017.0	0.9280	2014	0	0	0	0.0	0.0	1.0
15	Massachusetts General Hospital	2017.0	0.9280	2015	0	DISC	648	1.0	0.0	1.0
16	Massachusetts General Hospital	2017.0	0.9280	2016	0	HACK	4,293	1.0	0.0	1.0
17	Massachusetts General Hospital	2017.0	0.9280	2016	0	HACK	4293	1.0	0.0	1.0
18	Massachusetts General Hospital	2017.0	0.9280	2017	1	0	0	0.0	0.0	1.0
19	Massachusetts General Hospital	2017.0	0.9280	2018	0	DISC	588	1.0	0.0	1.0

Results

- Logit regression of pixel tracking usage indicator against privacy breach incident
- Hospital and year fixed effects

Results: Logit

```

=====
Model:                Logit                Pseudo R-squared:    0.553
Dependent Variable:   breach                AIC:                 639.4107
Date:                2023-04-20 23:04          BIC:                 741.6319
No. Observations:    3020                Log-Likelihood:      -302.71
Df Model:            16                LL-Null:             -676.61
Df Residuals:        3003              LLR p-value:         8.5454e-149
Converged:           0.0000              Scale:               1.0000
No. Iterations:      35.0000
=====

```

```

-----
                Coef.      Std.Err.      z      P>|z|      [0.025      0.975]
-----
pixel          1.2985      0.3770      3.4439  0.0006      0.5595      2.0375

```

Future plan of research

- Explore the outcome of pixel tracking usage
 - Social welfare
 - Con: discrimination
 - Pro: community quality of life
- Explore the determinants of pixel tracking usage
 - Centers for Medicare & Medicaid Services Data (CMS) dataset
 - Hospital profile (size, ownership, social responsibility, etc)
 - Cost report (Financial distress)
 - Market competition



The background of the slide features a large, faint watermark of the Rutgers University seal. The seal is circular with a sunburst in the center and the words "RUTGERS UNIVERSITY" and "THE STATE UNIVERSITY OF NEW JERSEY" around the perimeter.

RUTGERS

THE STATE UNIVERSITY
OF NEW JERSEY

The Benefits of Big Data: The Case of Wind Data to Detect Greenwashing

Yu Gu

Summer Project

What and Why

- Research Question:
 - Can wind data contain incremental information about the electricity generated by wind power from corresponding energy companies?
- Motivations:
 - Greenwashing is pervasive in the energy sector (RepRisk 2022)
 - Wind power is currently the second-largest renewable energy source globally for electricity generation (C2ES 2022)
 - The use of big data in accounting has gained attention, with researchers exploring diverse sources like scanner and satellite image data (Qian and Dichev, 2022; Gu, Dia, and Vasarhelyi, 2023)

How

- The quantity of electricity produced is a function of wind speed for a particular turbine (Lu and McElroy 2017)
- Wind data could provide evidence for
 - the quantity of electricity generated
 - the revenue produced
 - the environmental impact
- Further, detect the greenwashing practices, e.g., exaggerating the quantity of electricity produced by wind power on their ESG reports
- More importantly, investors could assess the company's performance in real-time and make informed decisions.

In Progress

- Collect data
 - Amazon Web Services (AWS)
 - The National Renewable Energy Laboratory (NREL) Wind Integration National Dataset
- Contact the NREL for collaboration



The background of the slide features a large, faint watermark of the Rutgers University seal. The seal is circular with a sunburst in the center and the words "RUTGERS THE STATE UNIVERSITY OF NEW JERSEY" around the perimeter.

RUTGERS

THE STATE UNIVERSITY
OF NEW JERSEY

Partner-client Social Ties and Audit Quality

Lanxin Jiang

Advised by Miklos Vasarhelyi

June 27, 2023

Motivation

- Tension
 - Embeddedness theory (Granovetter 1985) suggests that social networks can facilitate information transfer and alleviate management pressure to waive correction of detected misstatements.
 - However, social ties can also weaken monitoring mechanisms in governance (Guan, Su, Wu, and Yang 2016) and undermine auditors' monitoring of the financial reporting process (He, Pittman, Rui, and Wu 2017)
- Gap
 - Archival audit partner research largely uses non-US settings because the individual partner information was not available in the U.S. until 2017. Institutional differences can impede previous research findings from extending to the U.S. setting (Lennox and Wu 2018).

Objective and Contribution

- This research will explore different social ties between audit partners and client executives.
- This research will contribute to the auditing literature at the partner level in the U.S. setting, and literature using social media in the auditing setting.
- This research will add to the auditor social connection literature by examining different social ties and their impact on audit outcomes.

Progress

- ✓ Literature review
- Data collection
 - ✓ Audit partner names from PCAOB AuditorSearch Form AP database since 2017
 - ✓ Audit partner personal information from LinkedIn
 - ✓ Executive information from BoardEx
 - Audit outcomes (restatement, going concern opinions, discretionary accruals) from Compustat and Audit Analytics
- Construct social tie measures
 - employment affiliation connections (He et al. 2017)
 - school ties (Guan et al. 2016)
 - other social tie measures, e.g., membership, GNN centrality, link strength, neighborhood overlaps
- Examine their impact on audit outcomes and relative importance
- Additional analysis
 - Examine the interaction between partner's innate characteristics (e.g., gender,



The background of the slide features a large, faint watermark of the Rutgers University seal. The seal is circular with a sunburst in the center and the words "RUTGERS THE STATE UNIVERSITY OF NEW JERSEY" around the perimeter.

RUTGERS

THE STATE UNIVERSITY
OF NEW JERSEY

Governance and Donations: Evidence from Small Nonprofit Organizations

Katerina Liu, Rutgers University

GeDong, City University of New York

Chengzhang Wu, Marist College

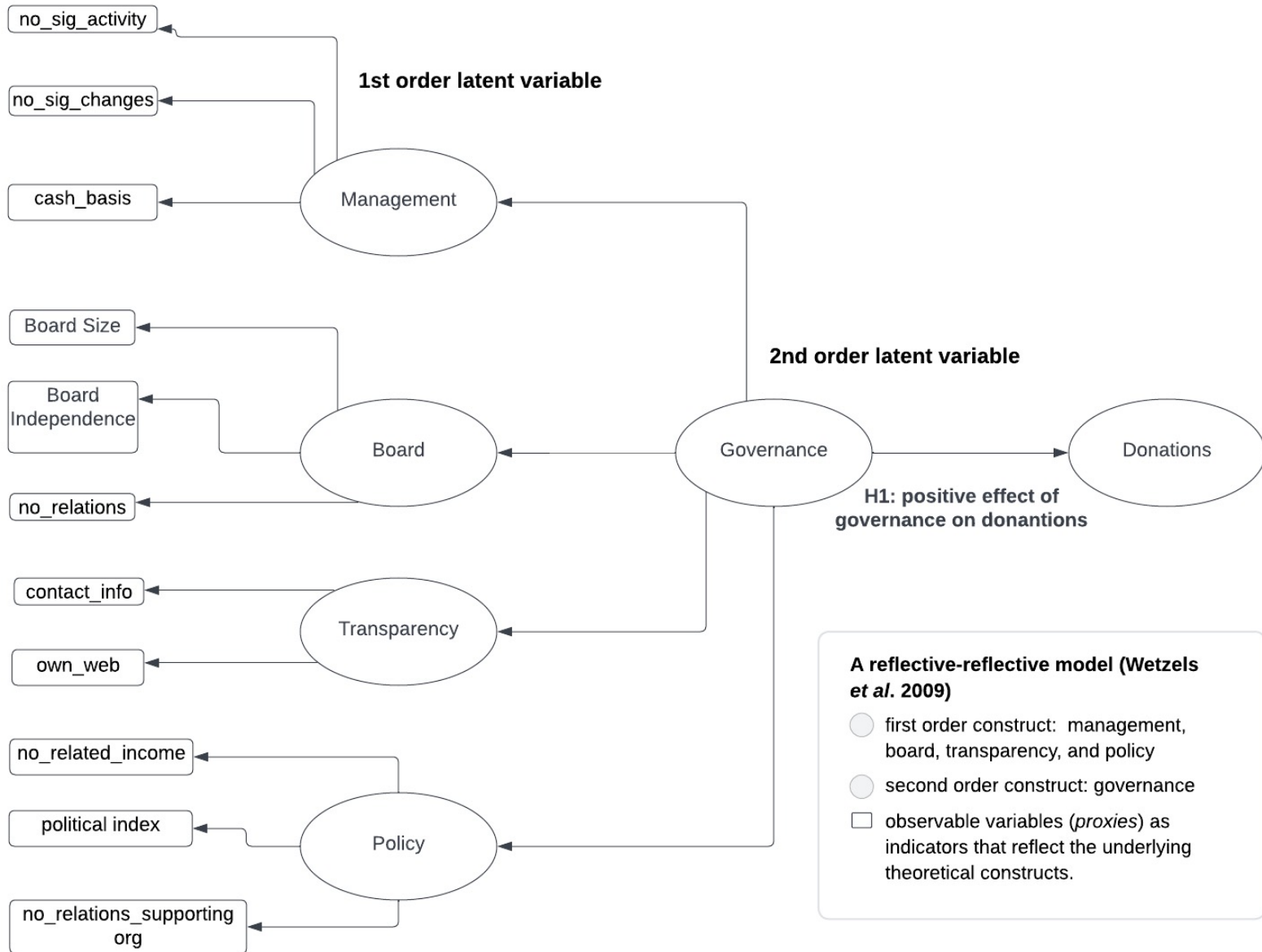
Contribution

- It is the first empirical study that examines the relation between governance and donations for Form 990-EZ filers.
- The development of the governance construct and the multiple conceptual layers of governance dimensions contributes to the agency theory approach in nonprofit literature.
- The operationalization of governance as a second-order construct makes it possible to measure the effect of governance on donations for Form 990-EZ filers.
- It sheds some light on regulatory oversight in the nonprofit sector for relevant stakeholders (Watchdog agencies and the general public).

Research Question and Key Findings

- Research Question: we study the governance and donations of Form 990-EZ filers.
- Key Findings
 - Among four governance constructs, three constructs (management, public access, and board) are positively associated with donations and one policy construct is negatively linked with donations.
 - Our results support the assertion that good governance disclosed by Form 990-EZ is value-added to donations that small charities receive, independent of other variables such as age, asset, wealth, program ratio, and fundraising expenditures.

Observed variable



2023

Tiny Totos – A case for proper ESG reporting

Jonas Sveistrup Søgaard, Ph.D.

Deloitte.
Digital



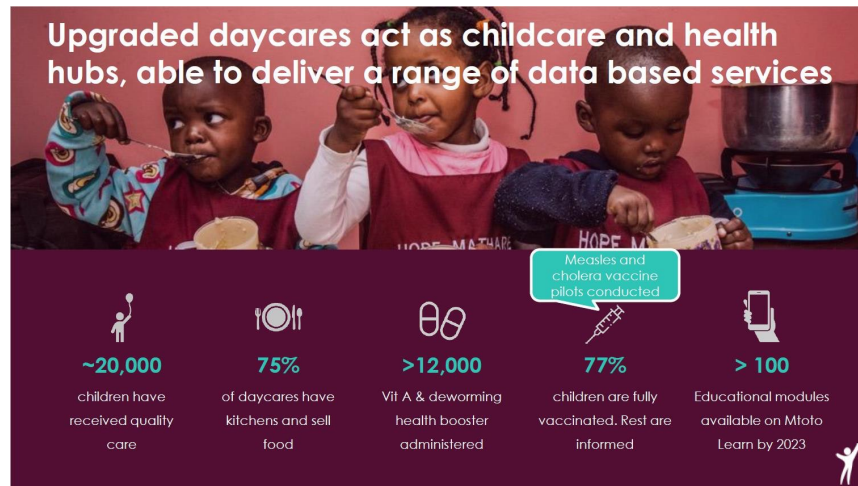
I seek co-author(s) for a case study for the social aspect of ESG reporting

I have access to data...

...I need co-author(s) who wants to develop the paper with me

I'm aiming for Special Issue Conference; Deadline 15th september

...however, not limited to that



Special Issue Conference
The Digital Transformation of ESG Reporting
Journal of Accounting and Public Policy and Journal of Information Systems

Increasingly stakeholders, such as investors, regulators/policy makers, and service providers are demanding greater reporting on environmental, social and governance (ESG) related matters. This enhanced need for ESG related information, combined with challenges associated with consistency and comparability in the required information, has reignited discussion of this topic and led to the drafting of new policies concerning ESG reporting. Herein despite advancements in information technology, and organizations increasing use of digitization of systems and processes to facilitate ESG reporting, factors such as data collection and data quality potentially affect ESG reporting. This creates a need to look at the intersection between ESG reporting and information technology.

In advancing our understanding of this critical issue, the *Journal of Accounting and Public Policy* and the *Journal of Information Systems* present this special issue as a collaboration between the two journals. Topics of interests include, but are not limited to:

- The impact of requirements for internal vs external information on ESG reporting.
- Digitization and digital transformation of systems and processes, and ESG reporting.
- Benefits or consequences of advances in information technology and their impact on ESG reporting.
- The impact of recent changes in ESG regulation on organizations' data collection and IT strategies.

All types of research design and methodologies are welcome. The conference will be jointly hosted by both journals and their publishers (Elsevier and the American Accounting Association) at Pace University (New York, US). Further details about the conference will be announced in the summer of 2023.

Important dates:
Submission due: Sep 15, 2023
1st round revision due: November 17, 2023
Special issue conference: December 1, 2023 (Friday) and December 2, 2023 (Saturday)

Journal publication:
The special issue is expected to be published in summer 2024. While submitting authors are encouraged to

Tiny Totos: A Kenyan social enterprise solving Africa's childcare crisis by turning babysitters into childcare entrepreneurs

We will give **1 million** children
and women in Africa access
to quality childcare by 2030

Tiny Totos is a Kenyan (TTK) social
enterprise working to solve
Africa's **childcare crisis**.

In Africa only **20% of mothers** can afford
quality, reliable childcare. **1 billion
children** will live in Africa by 2055. Poor
childcare doesn't just prevent women
becoming entrepreneurs; **poor childcare
keeps women and children trapped in
poverty for life**.

Tiny Totos has built a user-centred,
evidence-based, and tech-enabled

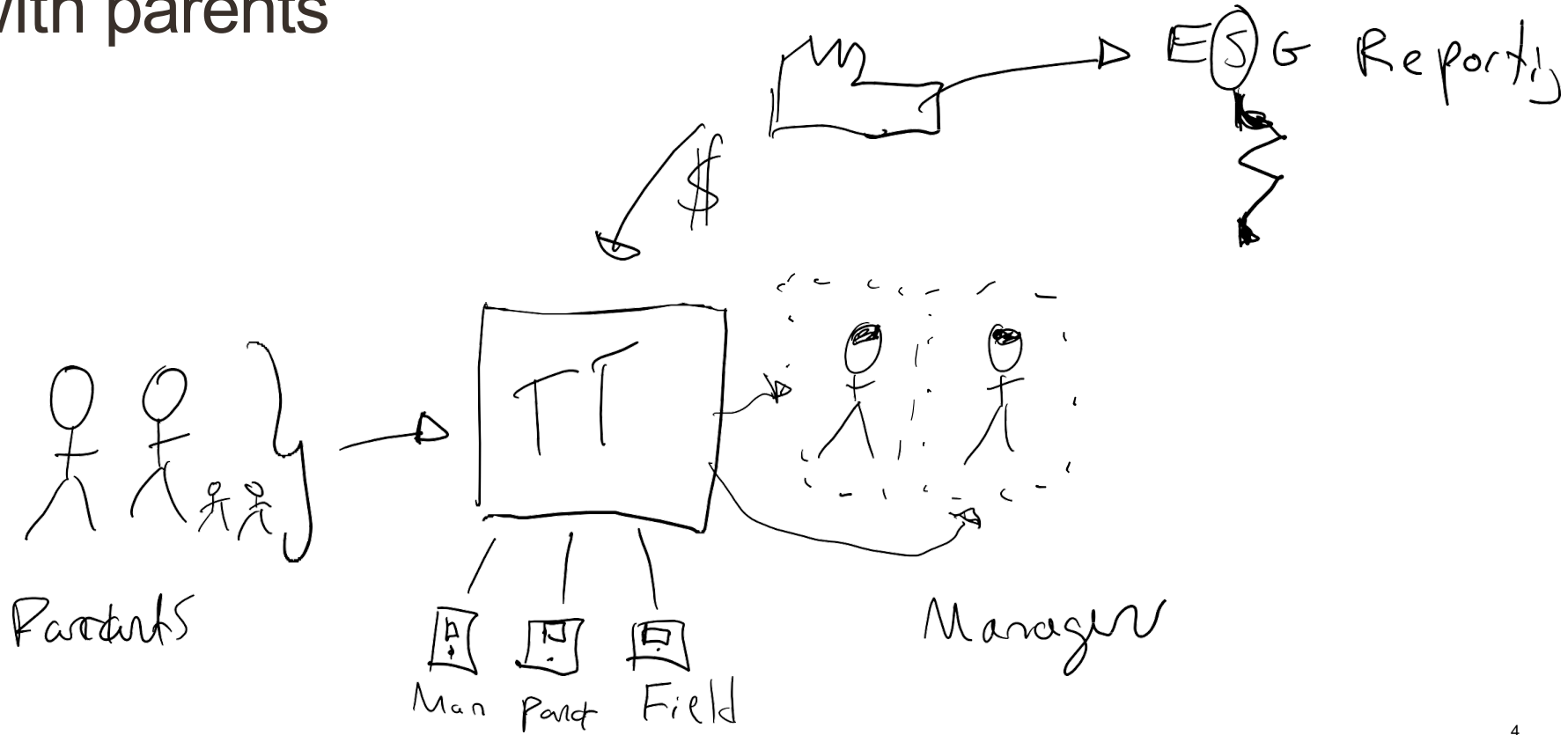
model for transforming quality and
sustainability of informal childcare.

When informal caregivers have access
to **network, capital, training and
technology**, quality of service improves.

**Babysitters become childcare
entrepreneurs**, earning decent
livelihoods without subsidy. Mothers can
focus on work without the handicap
of unreliable childcare. **Children in low
income settings** thrive, and look forward
to bright futures of their own.



Tiny Totos connects daycare managers with parents



The structure of the paper

1. Literature review in progress

Social Accounting – literature review

1 Social Accounting

Top cited papers with search terms "Social Accounting" or "ESG" in top accounting journals and top AIS.

Title	Year	Source title	Cited by
The effects of environmental, social and governance disclosures and performance on firm value: A review of the literature in accounting and finance	2018	British Accounting Review	264
The impact of environmental, social, and governance disclosure on firm value: The role of CEO power	2018	British Accounting Review	256
The effects of board characteristics and sustainable compensation policy on 2017 carbon performance of UK firms	2017	British Accounting Review	190
Social accounting's emancipatory potential: A Gramscian critique	2009	Critical Perspectives on Accounting	170
Voluntary environmental and social accounting disclosure practices in the Asia-Pacific region: An international empirical test of political economy theory	1999	International Journal of Accounting	158
Using neo-institutionalism to advance social and environmental accounting	2010	Critical Perspectives on Accounting	149
The impact of strategic relevance and assurance of sustainability indicators on investors' decisions	2015	Auditing	144
Cargo cult science and the death of politics: A critical review of social and environmental accounting research	2010	Critical Perspectives on Accounting	137
INVESTOR REACTION TO A CORPORATE SOCIAL ACCOUNTING	1984	Journal of Business Finance & Accounting	129
Twenty-one years of social and environmental accountability research: A coming of age	2011	Accounting Forum	118
ESG practices and the cost of debt: Evidence from EU countries	2021	Critical Perspectives on Accounting	105

2. Theory development not started

A couple of ideas

1. Extend recent paper from Gu, Dia and Vasarhelyi (2023)
2. ESG regulations to be used as the backbone (US and EU)
3. Propose a design for ensuring data on using blockchain as a trust infrastructure to ensure investor
4. Your input 😊

3. Data collection started

Two interviews have been conducted already. More are possible if needed, and access to their production data.

So; anyone up for a fun project to push the frontier of ESG reporting and help Tiny Totos ensure more funding by investors to help even more children in Africa?

Contact: jsveistrup@deloitte.dk or jss.acc@cbs.dk



The background of the slide features a large, faint watermark of the Rutgers University seal. The seal is circular with a sunburst in the center and the text 'RUTGERS THE STATE UNIVERSITY OF NEW JERSEY' around the perimeter.

RUTGERS

THE STATE UNIVERSITY
OF NEW JERSEY

The Increased Role of Advanced Technology and Automation in Audit: A Delphi Study

Research Focus

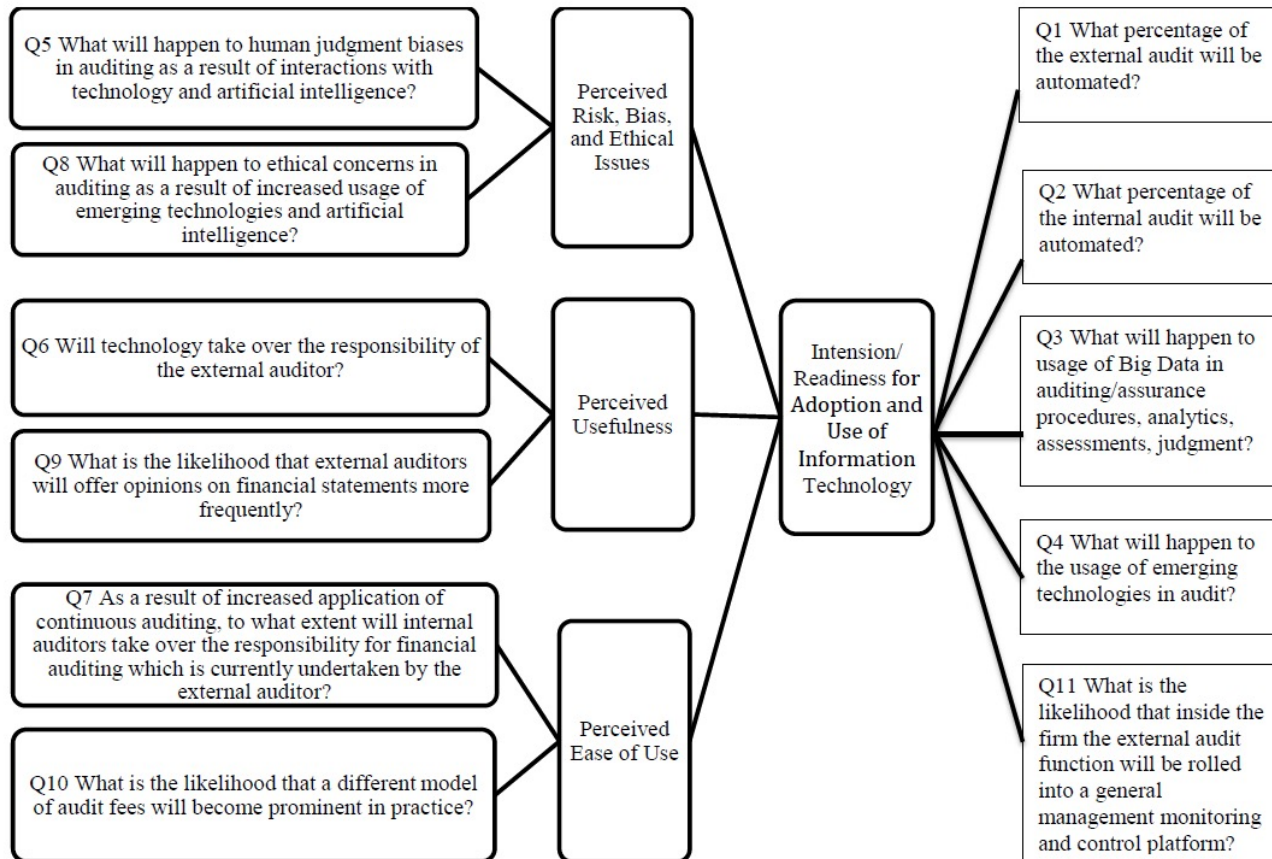
- What?
- This study seeks to gain expert predictions on the future directions of advanced technology and automation supporting auditors.

- Why?
- Covid-19: remote audit \implies accelerate the use of advanced technology
- Technological innovations \implies change the role of the accounting firm, staff, CPA

- How?
- Delphi Study (brainstorming session \implies Delphi)
- 19 experts

Research Method

Figure 1
Framework: Modified Technology Acceptance Model



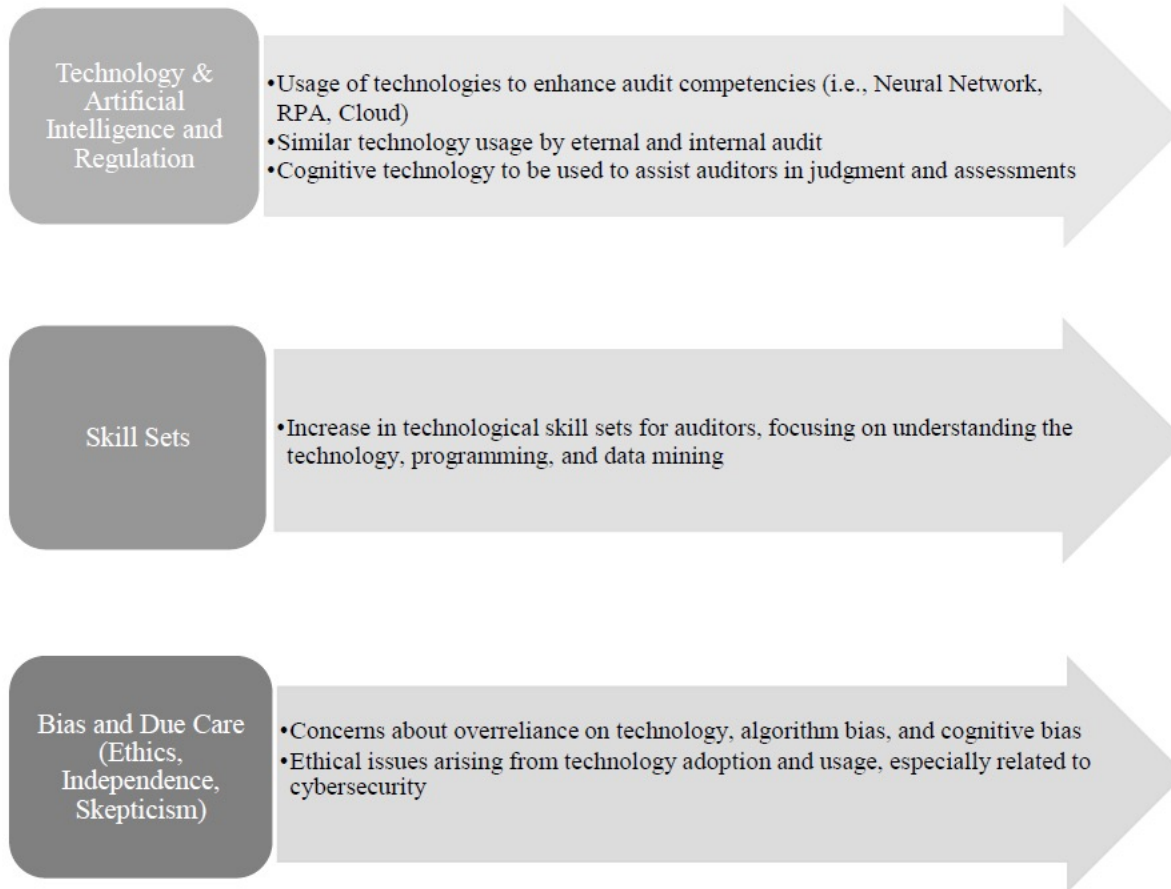
Finding

- Brainstorming session



Finding

Figure 2
Highlights, Predictions, and Recommendations Provided by Participants from the Brainstorming Session

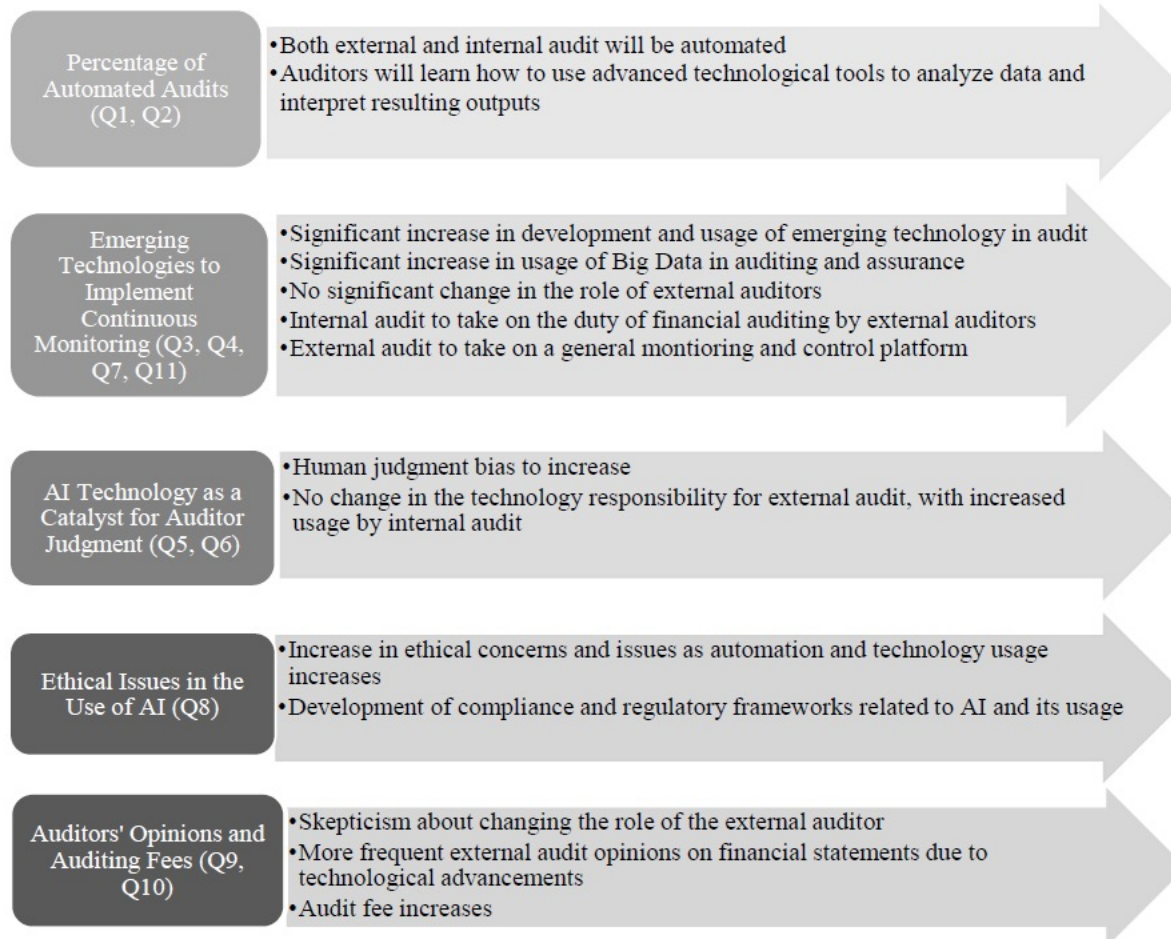


Finding

- Internal and external audits will be highly automated, with an increase in the usage of technology generally
- Technological skill sets of auditors are expected to greatly increase.
- There is concern about the ethical implications of using technology and the need for compliance and regulatory frameworks.

Finding

Figure 3
Highlights, Predictions, and Recommendations Provided by Participants from the Delphi Method



The background of the slide features a large, faint watermark of the Rutgers University seal. The seal is circular with a sunburst in the center and the words "RUTGERS THE STATE UNIVERSITY OF NEW JERSEY" around the perimeter.

RUTGERS

THE STATE UNIVERSITY
OF NEW JERSEY

Whether Firms' Risk Disclosures are Informative? Using a Novel Tool to Quantify Textual Data

Qingman Wu

Won Gyun No

Background

- Risk disclosures in corporate annual reporting are a very important information source for external information users to get more knowledge about the risks that firms face.

Using number of keywords

- Campbell et al. (2014) find that firms are intend to provide firm-specific and informative risk disclosures.
- Campbell et al. (2019) find that risk factor disclosures provide information about the level of a firm's future cash flows and investors incorporate this information into current stock prices.
- Chiu et al.(2019) find that risk factor disclosures provided by firms in their annual reports contain useful information that could potentially help their suppliers achieve better investment efficiency.
- Beatty et al. (2019) examine the change of information content of firms' risk disclosures before and after the financial crisis by separating risk into repeated risk, discontinued risk, new risk. They find firms' risk disclosures were informative before crisis and less informative after the crisis.

Using number of sentence related to firms' risks

- Kravet and Muslu (2013) examine the association between firms' risk disclosures and investors risk perception and find that firm-level risk disclosures are more likely to be boilerplate.
- Yang et al. (2018) examine the association between four risk measures derived from the risk factor section in 10-K filings and audit fees and find that firms' textual risk disclosures are informative.

Other methods

- Bao and Datta (2014) proposed a sent-LDA model (with 30 topics) to simultaneously discover and quantify risk types from textual risk disclosures. They find that some types of risks are informative and some are not.

Method

- Download 10-Ks from EDGAR
 - 10-Ks from 2005 to 2022 (tentative)
- Data preprocess
 - Parse out Item 1A from 10-K

ITEM 1A. RISK FACTORS

In addition to the other information in this report, the following risk factors should be considered before deciding to invest in any of Abbott's securities. Additional risks and uncertainties not presently known to Abbott, or risks Abbott currently considers immaterial, could also affect Abbott's business, financial condition, results of operations, or prospects could be materially adversely affected by any of these risks.

Business and Operational Risks

Abbott may acquire other businesses, license rights to technologies or products, form alliances, or dispose of or spin-off businesses, which could cause it to incur significant expenses and could negatively affect profitability.

Abbott may pursue acquisitions, licensing arrangements, and strategic alliances, or dispose of or spin-off some of its businesses, as part of its business strategy. Abbott may not complete these transactions in a timely manner, on a cost-effective basis, or at all, and may not realize the expected successful in making an acquisition, the products and technologies that are acquired may not be successful or may require significantly greater resources and investments than originally anticipated. Abbott may not be able to integrate acquisitions successfully into its existing business or transit efficiently, and could incur or assume significant debt and unknown or contingent liabilities. Abbott could also experience negative effects on its reported results of operations from acquisition or disposition-related charges, amortization of expenses related to intangibles and charges for impairments. These effects could cause a deterioration of Abbott's credit rating, result in increased borrowing costs and interest expense, and decrease liquidity.

Abbott depends on sophisticated information technology systems and maintains protected personal data, and a cyber attack or other breach affecting these information technology systems or protected data could have a material adverse effect on Abbott's results of operations.

Similar to other large multi-national companies, the size and complexity of the information technology systems on which Abbott relies for both its infrastructure and products makes them susceptible to a cyber attack, malicious intrusion, breakdown, destruction, loss of data privacy, or other. These systems have been and are expected to continue to be the target of malware and other cyber attacks. In addition, third party hacking attempts may cause Abbott's information technology systems and related products, protected data, or proprietary information to be compromised or stolen. Other disruption could result in adverse consequences, including increased costs and expenses, manufacturing challenges or disruption, problems with product functionality, damage to customer relations, reputational damage, lost revenue, and legal or regulatory penalties.

Abbott also collects, manages and processes protected personal data, including protected health information, in connection with certain medical products and service offerings. Abbott is subject to certain regional and local data protection laws that prohibit or restrict the transfer of protected personal information across international borders. For additional information concerning data privacy and security regulation, see the discussion in "Regulation" under Item 1, "Business." A breach of protected personal information could result in adverse consequences, including regulatory inquiries or litigation, increased costs and expenses, lost revenue, and fines or penalties.

Abbott invests in its systems and technology and in the protection of its products and data to reduce the risk of an attack or other significant disruption, and monitors its systems on an ongoing basis for any current or potential threats or vulnerabilities and for changes in technology and the regulatory environment. There can be no assurance that these measures and efforts will prevent future attacks or other significant disruptions to any of the systems on which Abbott relies or that related product issues will not arise in the future. Similarly, there can be no assurance that third party information technology breaches or other disruptions involving Abbott's systems or products could have a material adverse effect on Abbott's business.



The Watchdog Effect of Granular Transparency Regulation: Evidence from Firms' Responses to the Greenhouse Gas Reporting Program

Dong Gil Kim¹

¹Rutgers Business School

CARLAB meeting, 2023 June

Greenhouse Gas Reporting Program (GHGRP)

- The EPA implemented the Greenhouse Gas Reporting Program (GHGRP) in 2010, mandating facilities emitting over 25,000 metric tons of carbon dioxide equivalent per year to disclose their GHG emissions at the facility-level.
- Starting in 2012, the EPA began publicly releasing facility-level GHG emissions information through its website.
- Publicly traded companies owning regulated facilities are affected by the program's annual reporting requirements, as it offers granular transparency without imposing compulsory reduction measures.
- This unique regulatory context creates an opportunity to examine the influence of granular transparency regulation on firm-level behavior.

Relationship between firm and facilities

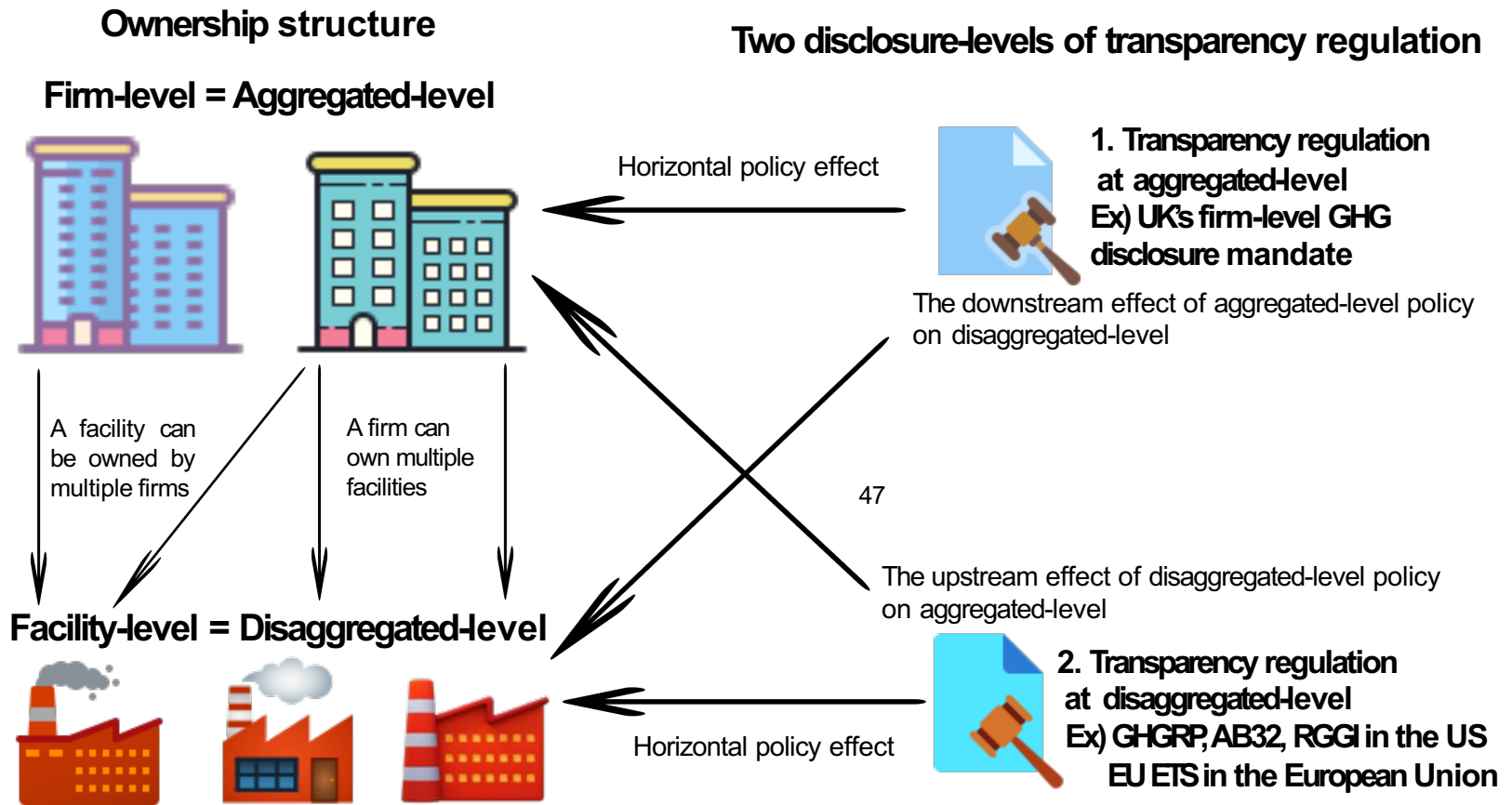
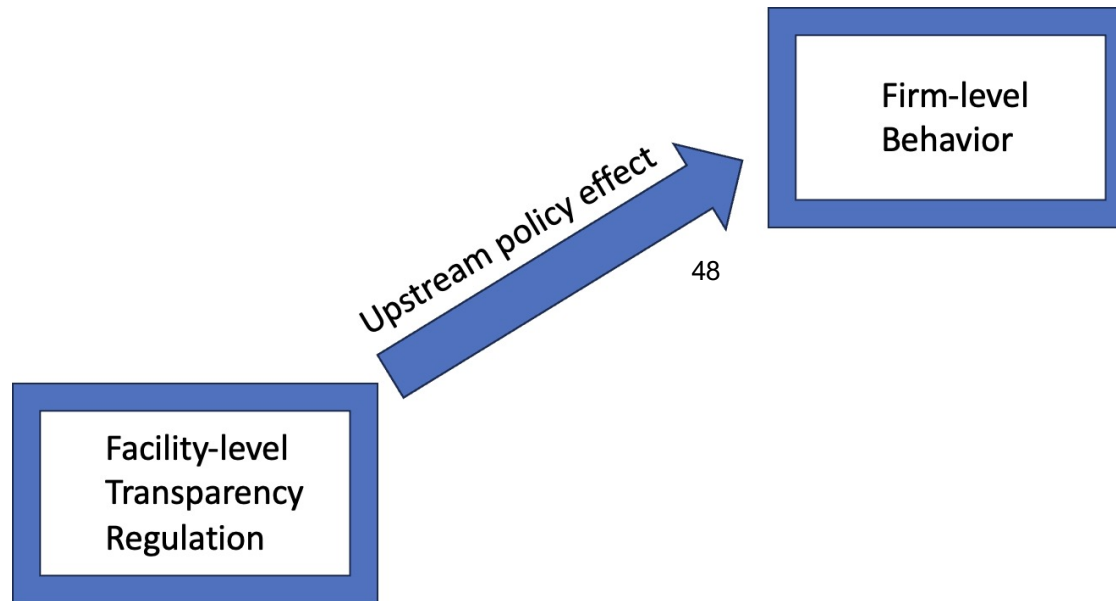


Figure: Firms, facilities and two disclosure-levels of transparency regulations

Research questions

- The key question is whether requiring facility-level GHG emissions disclosure has an influence on overall firm-level behavior.
- Specifically, does the mandate to disclose GHG emissions at the facility level prompt firms to voluntarily reduce their aggregate-level GHG emissions, and if so, what mechanisms are responsible for driving this effect?



Why do accounting students procrastinate? A qualitative analysis using chatgpt

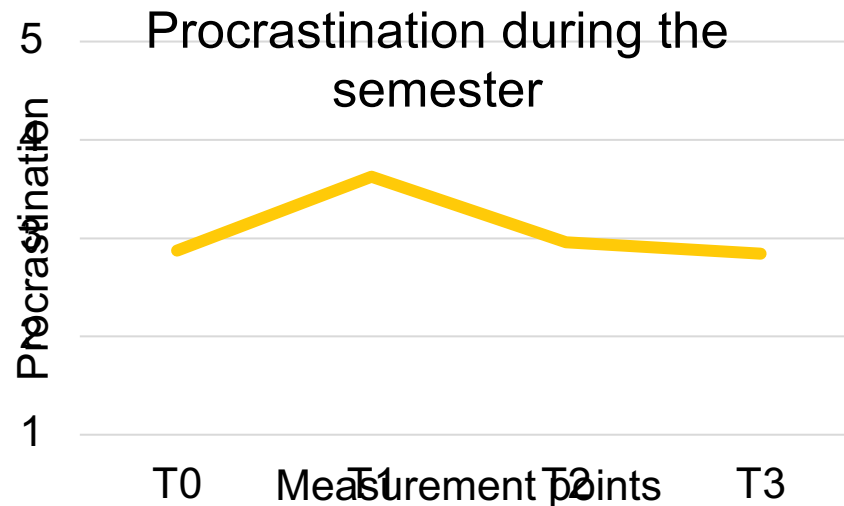
Eva Blondeel – Carlab Meeting 27-06-2023

Introduction

- Active learning interventions: effective to combat procrastination

(Blondeel et al., 2021; Blondeel et al., 2023a; Blondeel et al., 2023b)

- However, despite all efforts to actively engage students, the majority of the **students still procrastinate**



Source: Blondeel, Everaert & Opdecam (2023). Working paper.

Introduction

- Active learning interventions: effective to combat procrastination
(Blondeel et al., 2021; Blondeel et al., 2023a; Blondeel et al., 2023b)
- However, despite all efforts to actively engage students, the majority of the **students still procrastinate**

General RQ: Why do first-year accounting students procrastinate?

literature

Previous findings

- Self-regulation, self-discipline, impulsiveness (Steel & Klingsieck, 2016)
- Antecedents of procrastination (Klingsieck et al., 2013)
 - Personal antecedents: motivational, emotional etc.
 - Situational antecedents: social, external structure etc.
- Quantitative (Steel & Klingsieck, 2016) and qualitative (Klingsieck et al., 2013; Lao et al., 2023)

However ... there are differences in procrastination between **COURSES** (Kljajic et al., 2022)

- Accounting is perceived as a difficult course (Carnegie, 2021)
- Procrastination fluctuates throughout the semester of an accounting course (Blondeel et al., 2023 WP)

contribution

This study

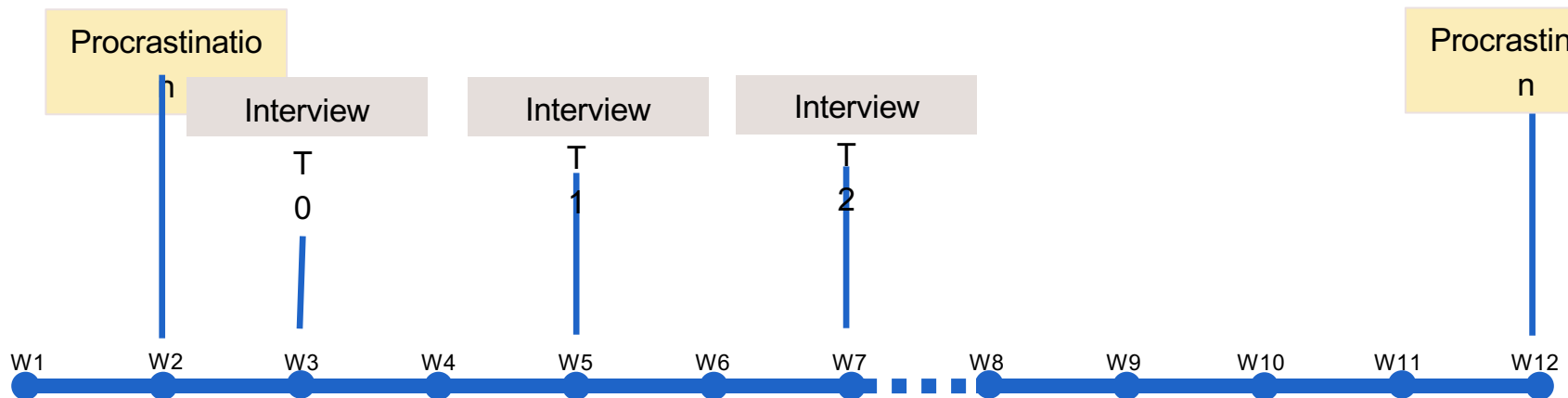
- Antecedents for procrastination in the accounting course
- Evolution of the antecedents throughout the semester

RQ1: Why are first-year accounting students procrastinating?

RQ2: How do these reason differ throughout the semester?

methodology

In-dept one-on-one interviews with first-year undergraduate accounting students in Business Economics at Ghent University



Qualitative data collection

3 Rounds of interviews

12 Different students each round

Methodology: sample

Stratified sample: Procrastination T0, Gender, and Ability

Round 1				Round 2				Round 3			
	Procrastination	Gender	Ability (/20)		Procrastination	Gender	Ability (/20)		Procrastination	Gender	Ability (/20)
S1	Medium	Male	11	S13	High	Female	8	S25	Low	Male	18
S2	High	Female	6	S14	High	Male	16	S26	High	Male	14
S3	High	Female	12	S15	Medium	Female	12	S27	High	Female	9
S4	Low	Female	15	S16	Medium	Male	16	S28	Medium	Female	10
S5	Medium	Female	14	S17	High	Female	15	S29	Medium	Male	19
S6	Medium	Male	13	S18	Medium	Female	16	S30	Medium	Female	bisser
S7	High	Female	9	S19	High	Female	14	S31	Medium	Male	13
S8	Medium	Female	7	S20	Low	Female	18	S32	Medium	Male	18
S9	High	Male	16	S21	High	Male	18	S33	High	Female	9
S10	Low	Male	12	S22	Medium	Female	6	S34	Low	Male	14
S11	Low	Female	16	S23	Medium	Male	15	S35	Medium	Female	8
S12	High	Female	6	S24	High	Female	15	S36	Medium	Female	13

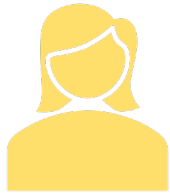
Methodology: approach

Qualitative interview data: two-step procedure

1. Transcribing the interviews
2. Dutch transcripts
3. Manual selection of sentences referring to procrastination

Methodology: approach

Qualitative interview data: two-step procedure



Manual coding

Develop coding scheme
Manual coding



ChatGPT coding

Provide coding scheme (or not ?)
Prompt Engineering (Gu et al., 2023)

Let's stay connected!

Eva Blondeel

Doctoral Researcher and Teaching Assistant

Research group Accounting Education, Ghent University



Eva.Blondeel@UGent.Be



[linkedin.com/in/eva-blondeel](https://www.linkedin.com/in/eva-blondeel)



www.accountingeducation.ugent.be



The background of the slide features a large, faint watermark of the Rutgers University seal. The seal is circular with a sunburst in the center and the text 'RUTGERS THE STATE UNIVERSITY OF NEW JERSEY' around the perimeter.

RUTGERS

THE STATE UNIVERSITY
OF NEW JERSEY

Artificial Intelligence Co-Piloted Auditing

Hanchi Gu, Marco Schreyer, Kevin Moffitt, and Miklos A. Vasarhelyi

Present by: Hanchi Gu

The Concept of Co-pilot



E.g., Tesla's Autopilot, Ford's BlueCruise, and Mercedes's Drive Pilot

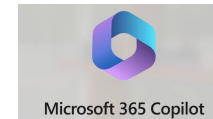
The Development of Co-pilot

Software Development
 (e.g., Visual Studio Code, GitHub Copilot)
 Customer Support
 (e.g., Zendesk, Intercom)

Healthcare
 (e.g., BuoyHealth, Ada Health)

Smart Home Assistants
 (e.g., Amazon Alexa, Apple Siri, Google Assistant),

General Productivity Assistant
 (Microsoft 365 Copilot)



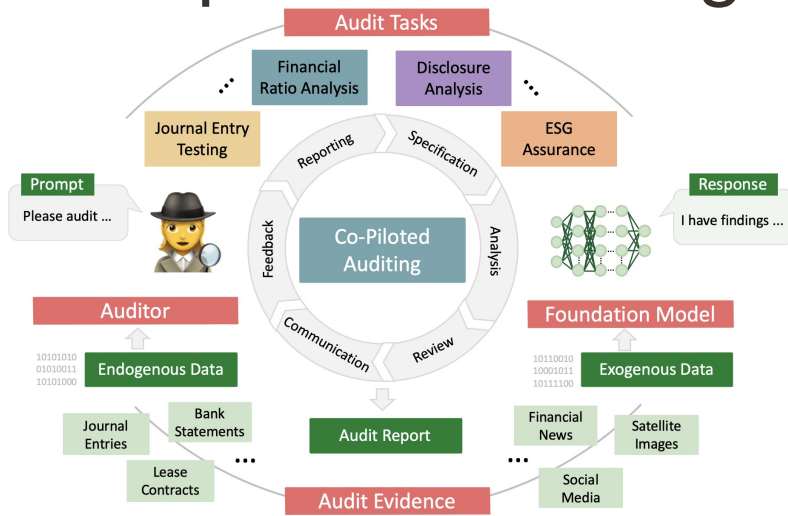
Auditing?

Summary

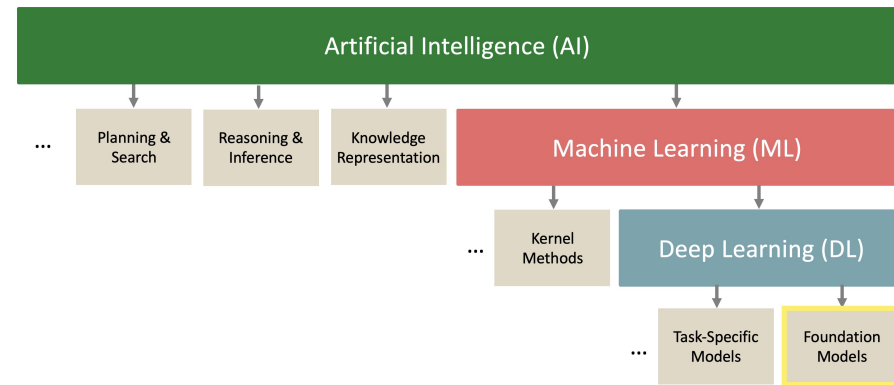
Our study offers the following contributions:

1. We introduce the concept of co-piloted auditing, a highly collaborative partnership between human auditors and AI models that capitalizes on each other's capabilities to enhance auditing.
2. We present a comprehensive overview of foundation models, with a focus on pre-training and fine-tuning techniques specifically tailored to the domain of financial auditing.
3. We demonstrate the effectiveness of employing chain-of-thought prompting to guide LLMs in addressing specific audit tasks, highlighting the model's versatility and reasoning capabilities.
4. We assess the effectiveness of our prompting protocols by employing ChatGPT, analyzing both the application of the protocols and the quality of the generated responses.

Foundation-Model-Empowered AI Co-piloted Auditing

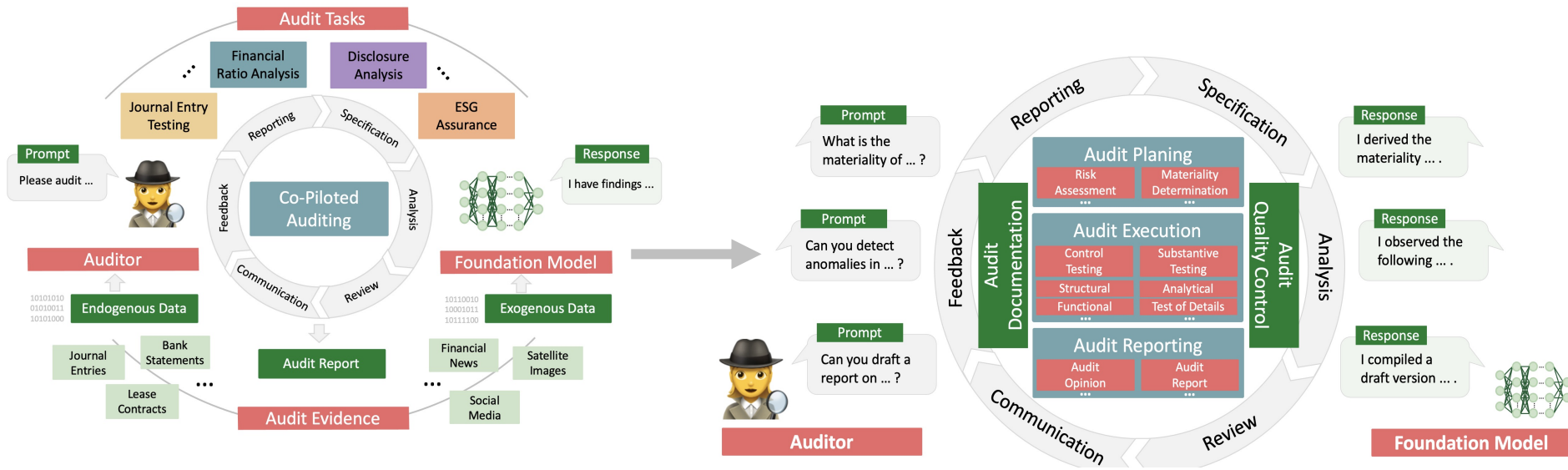


Co-piloted Auditing Concept Framework



Co-piloted Auditing Process

The Process of AI Co-piloted Auditing



Co-piloted Auditing Concept Framework

Co-piloted Auditing Process

Experiments of Audit Task

Experiment A: Audit Task -
Financial Ratio Analysis ✓

Experiment B: Audit Task
- Post-Implementation
Review ✓

Experiment C: Audit Task
Journal Entry Testing ✓

.....

Experiment A: Audit Task - Financial Ratio

Analysis

[Model]: Audit initiated. I am now calculating the three financial ratios based on the provided income statement data (Input 1).

Gross profit margin = Gross profit / Total revenue

Operating profit margin = EBIT / Total revenue

Net profit margin = Net income / Total revenue

Here are the calculated ratios:

Gross profit margin = $1.70782e+11 / 3.94328e+11 = 0.4330258827965754$

Operating profit margin = $1.19437e+11 / 3.94328e+11 = 0.3029328940914151$

Net profit margin = $9.9803e+10 / 3.94328e+11 = 0.25315486389767526$

So, the results are as follows:

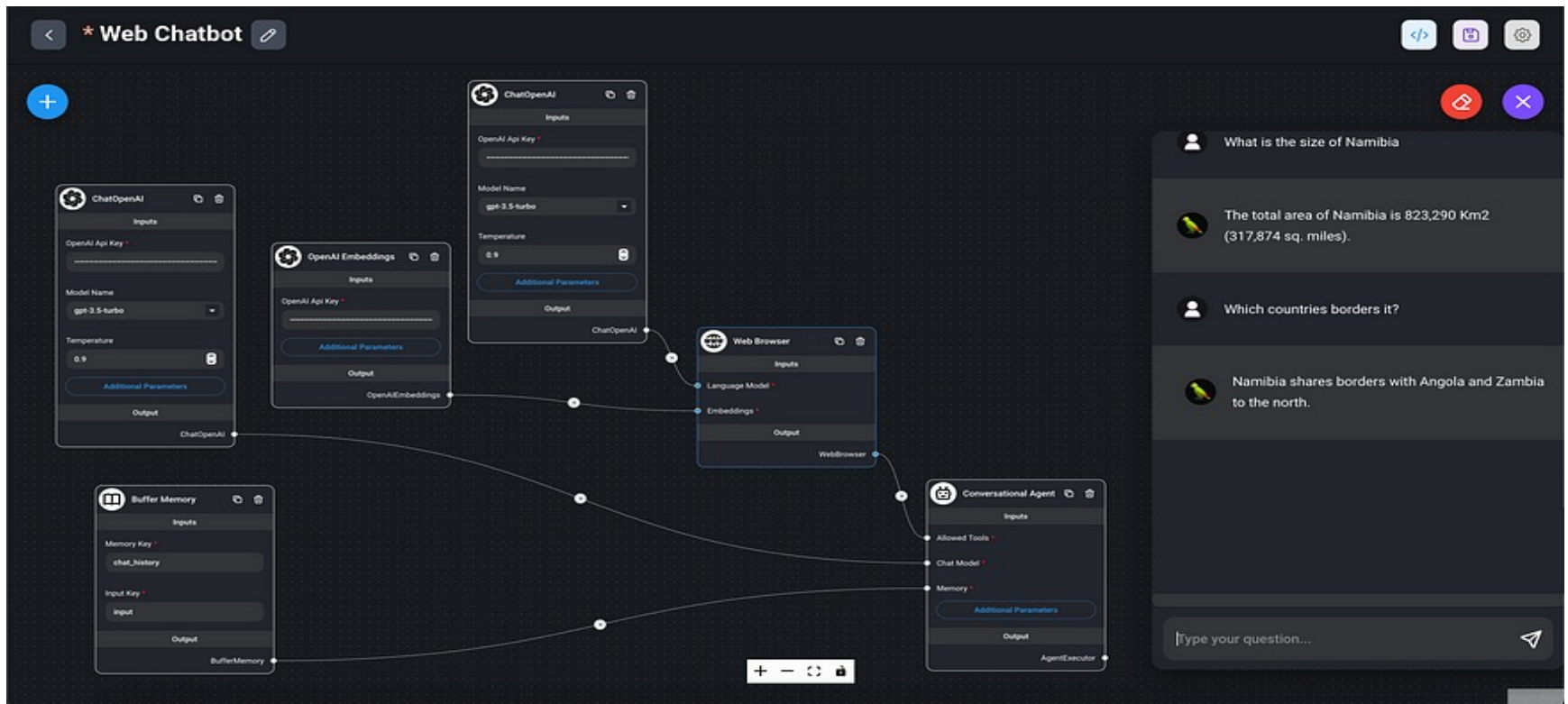
Calculation Completed!

Gross profit margin = 0.4330258827965754

Operating profit margin = 0.3029328940914151

Net profit margin = 0.25315486389767526

More Development



The background of the slide features a large, faint watermark of the Rutgers University seal. The seal is circular with a sunburst in the center and the words "RUTGERS UNIVERSITY" around the perimeter.

RUTGERS

THE STATE UNIVERSITY
OF NEW JERSEY

Manual Journal Entry Testing --Using Natural Language Processing and Deep Learning

Qing Huang, Huijue Kelly Duan, Marco Schreyer,
Miklos Vasarhelyi

Motivation

- Information flows within the organization automatically and seamlessly under ERP era.
- Flexibility to record complex transaction → Manual Journal Entries
- Current audit approach:
 - mainly relies on manual testing
 - randomly selecting, significant general ledger accounts, material dollar amounts
- Considering the volume of the MJE's that an organization needs to process, current sampling testing cannot facilitate an efficient and effective monitoring process

Methodology

- This study aims to find an innovative approach to evaluate MJE's and help the company's management to improve their control procedures over MJE's and detect anomalous or fraudulent MJE's.
- The study constructs three types of key risk indicators(KRI), including textual data and rare clusters, to identify potential fraudulent entries.
- Latent space embedding of KRI results are further clustered for downstream audit tasks
- The MJE's from a subsidiary of one multinational company are used to conduct this study. Three types of KRIs are constructed.

Key Risk Indicators

Rule-based

- Rule based analysis
 - Same User parking and posting
 - Missing key information
 - Materiality
 - Round number
 - Certain reclassification entries (Sales/Marketing ↔ Discount, etc.)

Transaction Anomaly

Deep Learning

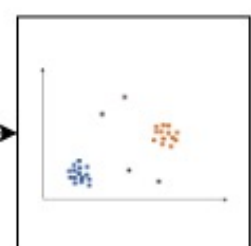
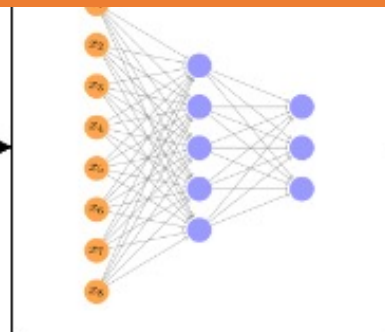
Textual Anomaly

Latent Space Representation

Clustering

	v_1	v_2	v_3	v_4	...	v_k
j_1	5200005	2022	12	A	...	SA
j_2	5200007	2022	11	D	...	DR
j_3	5200012	2022	12	C	...	DR
j_4	5800068	2021	2	B	...	DR
j_5	6100075	2022	12	C	...	DR
j_6	6100121	2022	11	C	...	AR
j_7	6900087	2022	11	A	...	AR
j_8	7000196	2022	11	B	...	AR
...	AR
j_n	9800031	2022	11	A	...	AR

Journal Entries



Contribution

- Presents an innovative way to comprehensively and systematically evaluate JEs, and assist the company's management in improving its control procedures over JEs.
- Enhances the company's control and monitoring procedures over MJEs.
- Textual analysis in this study can provide management more intuitive information regarding the transaction essences.
- Latent space representation of KRI yield interpretable anomaly transaction identification



The background of the slide features a large, faint watermark of the Rutgers University seal. The seal is circular with a sunburst in the center and the words "RUTGERS UNIVERSITY" and "THE STATE UNIVERSITY OF NEW JERSEY" around the perimeter.

RUTGERS

THE STATE UNIVERSITY
OF NEW JERSEY

Collaborate Verbal, Voice, and Visual Evidence in Detecting Managerial Uncertainty and Financial Outcomes

Yue Liu

Second Year Summer Research

RUTGERS

THE STATE UNIVERSITY
OF NEW JERSEY

Research Question:

Whether the uncertainty evidence in the interviews of CEO about earnings can be used to predict financial outcomes?

Contribution:

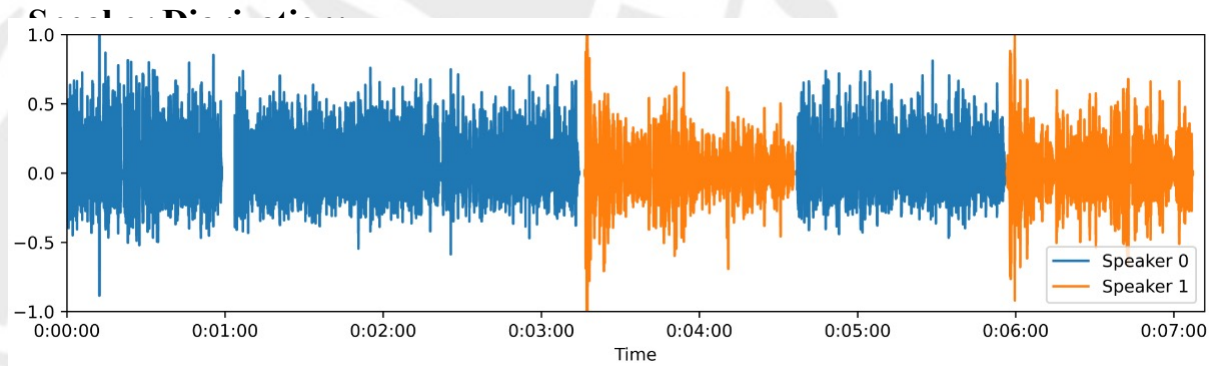
- Unstructured evidence
- Generalize laboratory conclusions to real business situations

Behavior experiment in accounting(Cade et al., 2020):

Investors react more negatively to a video disclosure when the CEO displays nonverbal cues of uncertainty.

Sample Selection:

- 2018.01.01 — 2023.05.31
- 545 videos
- Keywords searching and manual processing



Extract Frames From Video:
Frames per second

RUTGERS

THE STATE UNIVERSITY
OF NEW JERSEY

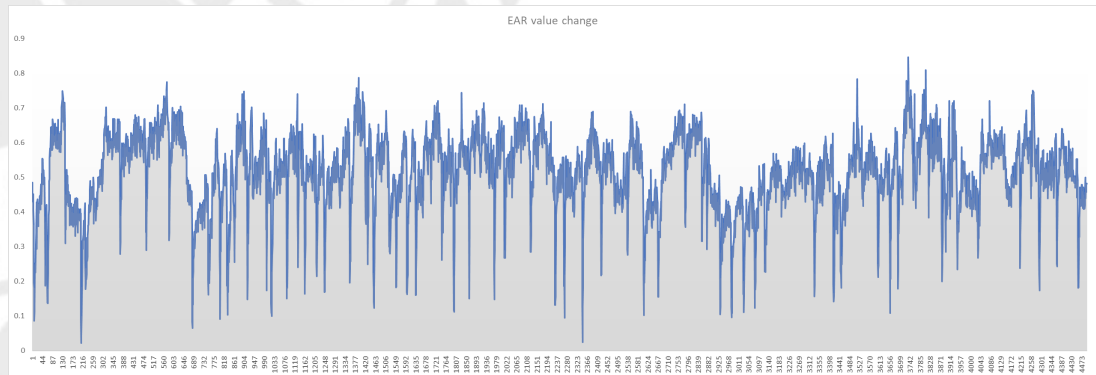
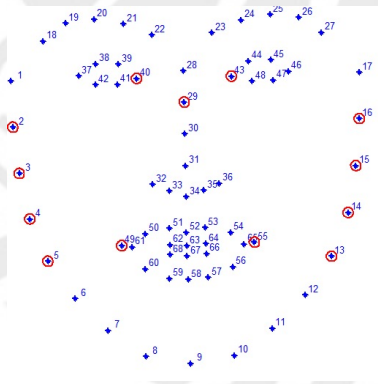
Voice Evidence:

- Pitch
- Harmonics-to-Noise Ratio
- Jitter and Shimmer
- ...

Measure Duration	Measure Pitch	Measure Harmonics-to-Noise-Ratio	Measure Jitter	Measure Shimmer	Measure Cepstral Peak Prominance (CPP)	Measure Formants	Measure Vocal Tract Estimates	Measure Intensity
	Measure Pitch							
Pitch Praat To Pitch (ac)								
Mean Pitch (F0) (Praat To Pitch (ac))	180.04600838006442							
Median Pitch (F0) (Praat To Pitch (ac))	171.8613868897539							
Standard Deviation Pitch (F0) (Praat To Pitch (ac))	57.57692059810071							
Pitch Min (F0) (Praat To Pitch (ac))	96.79468309675899							
Pitch Max (F0) (Praat To Pitch (ac))	500.2397148204961							
Pitch Floor	100							
Pitch Ceiling	500							
Pitch To Pitch (cc)								
Pitch Values (Praat To Pitch (cc))	[215.99860091878196, 215.25574682812226, 215.5286970353256, 215.58038915131004, 213.58595193728956, 213.57455641191166, 209.984943341							

Facial Evidence:

- Expression
- Eye movement
- Frowning
- Eyebrows Raising
- Head Side Turn
- ...



Can External Auditors Utilize ChatGPT to Assure ESG Reports? - Starting from Environmental Section Test

Presenter: *Irina (Jie Bao)*

Coauthored with *Benita Gullkvist*

6.27.2023

Research goals:

- ✓ Explores the potential and limitations of ChatGPT in reviewing ESG reports.
- 1. Design a Framework:
 - to show how we could utilize ChatGPT to extract information and further help assure ESG reports (an example of environmental section tests)
- 2. Empirically Test:
 - Whether ChatGPT's outputs agree with the manual check results?

Background

- ESG/ sustainability reports and its assurance
- AI implementation in the Audit process
- Emergence of OpenAI and GPT Architecture
 - Strengths of ChatGPT
 - Opportunities & Benefit v.s. Challenges & limitations

Assurance Process

1. Review and understand the company's ESG reporting framework:
 - Global Reporting Initiative (GRI)
 - Sustainability Accounting Standards Board (SASB)
 - Task Force on Climate-related Financial Disclosures (TCFD)
2. Assess the company's data collection and reporting processes
3. Verify the accuracy of reported emissions data: Assess the company's emissions reduction targets and performance:
4. Review disclosures and reporting
5. Provide assurance opinion

Assurance Process

1. Review and understand the company's ESG reporting framework:
 - **Global Reporting Initiative (GRI) - GRI 305: Emissions**
 - Sustainability Accounting Standards Board (SASB)
 - Task Force on Climate-related Financial Disclosures (TCFD)
2. Assess the company's data collection and reporting processes
3. Verify the accuracy of reported emissions data: Assess the company's emissions reduction targets and performance:
4. **Review disclosures and reporting**
5. Provide assurance opinion

ESG Reporting sample (partial):

Scope 1 and Scope 2 GHG Emissions (MT CO2e)

	2019	2020	2021
Scope 1	9,192	10,961	11,447
CO2	7,693	9,380	9,561
CH4	239	252	301
N2O	1,260	1,329	1,585
Scope 2 location-based	39,735	41,705	36,577
Total Scope 1 + Scope 2	48,927	52,666	48,024

Note: The 2021 increase in Scope 1 emissions is due to the addition of the Frederick, Colorado and Lexington, Massachusetts sites.

Scope 3 Emissions (MT CO2e)

	2019	2020	2021
Agilent Fleet (US)	8,849	6,911	6,976
Transportation and Distribution*	139,715	101,453	68,753
Business Travel**	20,437	6,906	3,640
Total Scope 3 Emissions	169,001	115,270	79,369

* Kintetsu World Express, one of our third-party logistics vendors, reported a significant reduction of estimated CO2e emissions in 2021 based on a change in the emissions calculations methodology for long- and short-haul flight factors

** Business Travel: Air & Rail, Rental Cars

GHG Emissions Intensity

	2019	2020	2021
Scope 1 + 2 Intensity (MT CO2e/ft2)	0.01283	0.01320	0.01124

Reduction Of GHG Emissions GRI from new energy projects

	2019	2020	2021
Scope 1 + 2 (MT CO2e)	492	2795	970

Agilent_2021

GRI 305: Emissions

Topic disclosures:

1. Disclosure 305-1 Direct (Scope 1) GHG emissions
2. Disclosure 305-2 Energy indirect (Scope 2) GHG emissions
3. Disclosure 305-3 Other indirect (Scope 3) GHG emissions
4. Disclosure 305-4 GHG emissions intensity
5. Disclosure 305-5 Reduction of GHG emissions
6. Disclosure 305-6 Emissions of ozone-depleting substances (ODS)
7. Disclosure 305-7 Nitrogen oxides (NO_x), sulfur oxides (SO_x), and other significant air emissions

- The reporting organization **shall report the following information:**
 - a. **Gross direct (Scope 1) GHG emissions in metric tons of CO2 equivalent.**
 - b. **Gases included in the calculation; whether CO₂, CH₄, N₂O, HFCs, PFCs, SF₆, NF₃, or all.**
 - c. **Biogenic CO2 emissions in metric tons of CO₂ equivalent.**
 - d. **Base year for the calculation, if applicable, including:**
 - i. the rationale for choosing it;
 - ii. emissions in the base year;
 - iii. the context for any significant changes in emissions that triggered recalculations of base year emissions.
 - e. **Source of the emission factors and the global warming potential (GWP) rates used, or a reference to the GWP source.**
 - f. **Consolidation approach for emissions; whether equity share, financial control, or operational control.**
 - g. **Standards, methodologies, assumptions, and/or calculation tools used.**

Compilation requirements

Questions to test:

- Q1: “How much are scope 1 GHG emissions in metric tons of CO₂e?”
- Q2: “How much are scope 2 GHG emissions in metric tons of CO₂e using the location-based calculation methodology?”
- Q3: “How much are scope 2 GHG emissions in metric tons of CO₂e using the market-based calculation methodology?”
- Q4: “How much are scope 3 GHG emissions in metric tons of CO₂e?”
- Q5: “How much are the biogenic CO₂ emissions in metric tons of CO₂ equivalent?”
- Q6: “What gases (CO₂ /CO₂ , CH₄ /CH₄ , N₂O, HFCs, PFCs, SF₆ , NF₃) are included in the GHG emission calculation?”
- Q7: “What is the base year (baseline) for the emission calculation disclosed?”
- Q8: “What is the rationale for choosing the base year (baseline)?”
- Q9: “How much are scope 1 GHG emissions in metric tons of CO₂e in the base year (baseline)?”

Questions to test:

Q10: “What significant changes in emissions triggered recalculations of base year emissions?”

Q11: “What are the sources of the emission factors?”

Q12: “What are the global warming potential (GWP) rates?”

Q13: “What reference is used for the GWP source?”




Q14: “What is consolidation approach for GHG emissions?; Is it equity share, financial control, or operational control?”

Q15: “What are the standards, methodologies, assumptions, or calculation tools used in the GHG emissions calculation?”

Q16: “Does the calculation of gross direct (Scope 1) GHG emissions exclude any GHG trades?”

Q17: “What reporting standard/protocol/framework did the file use?” Is it GRI, SASB, TCFD, ISO, CDP, WRI?”

Method

- Feed the 17 questions to ChatGPT and record the outputs
- Methods tried:
 - Plug-in
 -  AskYourPDF
 -  WebPilot
 -  Link Reader
 - API

Sample code to use ChatGPT API

```

#Ask ChatGPT
num = 0
tlist1 = tqdm.tqdm(document)
for news in tlist1:
    #try:
        response = openai.ChatCompletion.create(
            model=model_engine,
            messages=[
                {"role": "user", "content": "Tell me within 5 words what is the the Government level (answer by Federal=1, state=2, local=3, other=4) of the position:" + news},
                {"role": "user", "content": "Explain why within 20 words"},
            ],
            max_tokens=40,
            n=1,
            stop=None,
            temperature=0,
        )
        raw = response.choices[0]['message']['content'].strip()
        answers.append(raw)
        num+=1
    #except:
        #raw = "Error"
        #answers.append(raw)
        #num+=1

df_save = df
df_save['Answers'] = answers

ques = {}
for i in range(14):
    n=i+1
    ques["Q" + str(n)] = []
    ques['Title'] = []

for i in df_save['Answer']:
    seperate = i.split(' A')
    for i in range(14):
        n=i+1
        ques["Q" + str(n)].append(seperate[i])
  
```


Method

Sample:

- use the year 2021' ESG reports
- S&P 500 companies without financials (SIC6) = 400 companies
- about 260-280 use GRI standards
- Further check the ones with separate Auditor's assurance reports



- ✓ Current sample: 54 (recorded all manual checked answers to 17 questions)
- ✓ Whether ChatGPT's outputs agree with the manual check results

Method

- Code the answers as 1 if it has found the answer, and 0 if it's N/A
- Check Type1 and Type2 errors

		Manual Check(correct one)	
		possitive(0)	Negative(1)
ChatGPT	possitive(0)	TP	FP
	Negative(1)	FN	TN

- Could further infer the Reporting Quality via the number of questions answered (i.e. contents disclosed)



The background of the slide features a large, faint watermark of the Rutgers University seal. The seal is circular and contains the text 'THE STATE UNIVERSITY OF NEW JERSEY' around its perimeter. In the center of the seal is a sunburst design with rays extending outwards.

RUTGERS

THE STATE UNIVERSITY
OF NEW JERSEY

Sentimental Analysis as a measure for determining the Informativeness of Cybersecurity Disclosures

Presented by: Hongmin William Du

Introduction

- Importance of Cybersecurity Disclosures
 - Definition of cybersecurity disclosures
 - Significance of accurate and informative disclosures for stakeholders, investors, customers, and employees
 - Impact of cybersecurity disclosures on decision-making, risk assessment, and overall trust in organizations
 - Challenges in evaluating the informativeness of cybersecurity disclosures using traditional measures
- Application of Sentimental Analysis on Cybersecurity Disclosures
 - Definition of sentimental analysis
 - Advantages of sentimental analysis compared to traditional methods

Research Question and Hypothesis

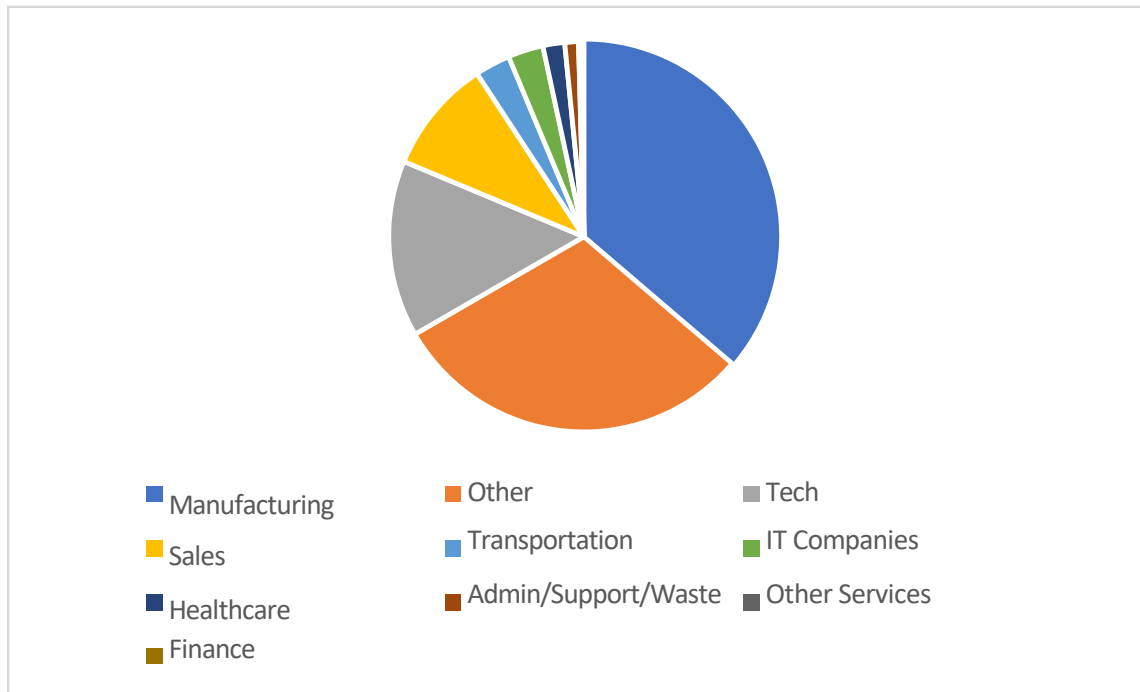
- Research Question: Can sentimental analysis be a better measure in determining the informativeness of cybersecurity disclosure?
- Hypothesis: Traditional measures of cybersecurity risk, without consideration of sentiment, can be misleading, as firms with more detailed cybersecurity disclosures may actually have lower actual cybersecurity risk.

Data and Methodology

- Five main sets of data
 - Firm Cybersecurity Disclosures (10-K)
 - Unique dataset of software's used by firms (Stackshare)
 - Software vulnerabilities (NIST Vulnerability Database)
 - Standard firm control variables (Compustat)
 - Cybersecurity Breach Dataset (PRC)
- Main unique dependent variable: Firm average software vulnerability score
- Sub- dependent variable (traditional): Firm data breach (0 if no breach, 1 if breach)
- Independent variable: Firm cybersecurity disclosure sentiment score

Sample Distribution

- Final Sample: 464 unique firms
- Spans a variety of industries



Regression and Results

$$\log(\text{real_risk_score}_{it}) = \beta \times \text{sentiment_score}_{it} + X_{i,t-1} \quad (2)$$

Table VIII. Informativeness of Sentiment Score of Real Cybersecurity Risk

Table reports regression results for the predictability of a firm's real cybersecurity risk using the sentiment of their cybersecurity disclosure as specified in Equation (2). The y-variable in all of the specifications is the real cybersecurity risk of a firm i in year t (construction of measure is described in Section 4). The coefficient of interest is the *sentiment_score* which is the sentiment of a firm's cybersecurity disclosure. Lagged firm control variables include a firm's profitability (*prof*), tangibility (*tangl*), log(size) (*logal*), market-to-book ratio (*mk2bkl*), and market leverage (*ml*). All regressions are double clustered by firm and year.

	(1)	(2)	(3)	(4)	(5)	(6)
	log(real_risk_score)	log(real_risk_score)	log(real_risk_score)	log(real_risk_score)	log(real_risk_score)	log(real_risk_score)
sentiment_score	-0.14** (0.036)	-0.15** (0.040)	-0.12* (0.040)	-0.090* (0.037)	-0.087 (0.038)	-0.020 (0.011)
prof		0.062 (0.14)	-0.0100 (0.13)	-0.021 (0.14)	-0.031 (0.14)	-0.0027 (0.099)
tangl		-1.19 (1.07)	-0.41 (1.03)	-0.58 (0.99)	-0.42 (0.99)	0.91 (0.65)
logal		0.26** (0.077)	0.31** (0.077)	0.27** (0.080)	0.26** (0.080)	0.18 (0.11)
mk2bkl		0.0090 (0.0059)	0.0085 (0.0061)	0.0062 (0.0047)	0.0056 (0.0050)	0.000020 (0.0025)
ml		-1.41 (0.76)	-0.64 (0.73)	-0.56 (0.74)	-0.57 (0.75)	-0.24 (0.27)
Constant	13.9*** (0.39)	12.4*** (0.77)	11.8*** (0.86)	12.3*** (0.67)	12.3*** (0.67)	12.8*** (0.77)
Industry FE	No	No	Yes	Yes	No	No
Firm FE	No	No	No	No	No	Yes
Year FE	No	No	No	Yes	No	Yes
Ind-Year FE	No	No	No	No	Yes	No
FirmYear Cluster	Yes	Yes	Yes	Yes	Yes	Yes
N	1462	1385	1304	1304	1288	1330
R2	0.017	0.056	0.22	0.26	0.20	0.94

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

Conclusion

- Sentiment Analysis shows promise as a valuable tool in evaluating the informativeness of cybersecurity disclosures
- Sentiment Analysis allows for a more comprehensive understanding by capturing subjective information
- Improved decision making: By considering sentiment analysis, organization can make more informed decisions, respond to stakeholders' concerns, and proactively address cybersecurity risks
- Continued research is necessary to refine sentiment analysis methodologies, and establish best practices for its implementation



The background of the slide features a large, faint watermark of the Rutgers University seal. The seal is circular with a sunburst in the center and the words "RUTGERS UNIVERSITY" around the perimeter.

RUTGERS

THE STATE UNIVERSITY
OF NEW JERSEY

Summer Research Progress Presentation

Analyzing Analyst Question Patterns in Earnings Calls: Leveraging NLP Models for Sentiment and Forward-Looking Analysis

Haoyun (Harry) Gao

06.27.2023

Analysts in conference call Q&A sections

- seems random

Q2 2023 Apple Inc Earnings Call - Final

Questions and Answers

OPERATOR: Certainly. We will go ahead and take our first question from Erik Woodring of Morgan Stanley.

ERIK WILLIAM RICHARD WOODRING, RESEARCH ASSOCIATE, MORGAN STANLEY,

RESEARCH DIVISION: Tim, maybe if we start with you. If we go back to the December quarter and the shutdowns -- the production shutdowns around the time, I think the question a lot of us were asking was, how should we think about the deferral of demand versus destruction of demand? March quarter was quite strong, 2% year-over-year iPhone growth.

Q1 2023 Tesla Inc Earnings Call - Final

Questions and Answers

MARTIN VIECHA: Thank you very much. And let's go to investor questions on say.com. The first one is what is the process to make auto pricing adjustments? What variables do you consider? How frequently do you review pricing?

MARTIN VIECHA: Great. Thank you very much. And let's go to analyst questions. We'll start with Alex Potter from Piper Sandler.

ALEXANDER EUGENE POTTER, MD & SENIOR RESEARCH ANALYST, PIPER SANDLER & CO., RESEARCH DIVISION: Can you hear me?

MARTIN VIECHA: Yes.

ELON R. MUSK: Yes.

ALEXANDER EUGENE POTTER: Okay. Perfect. So first question was on Lathrop.

In a Nutshell

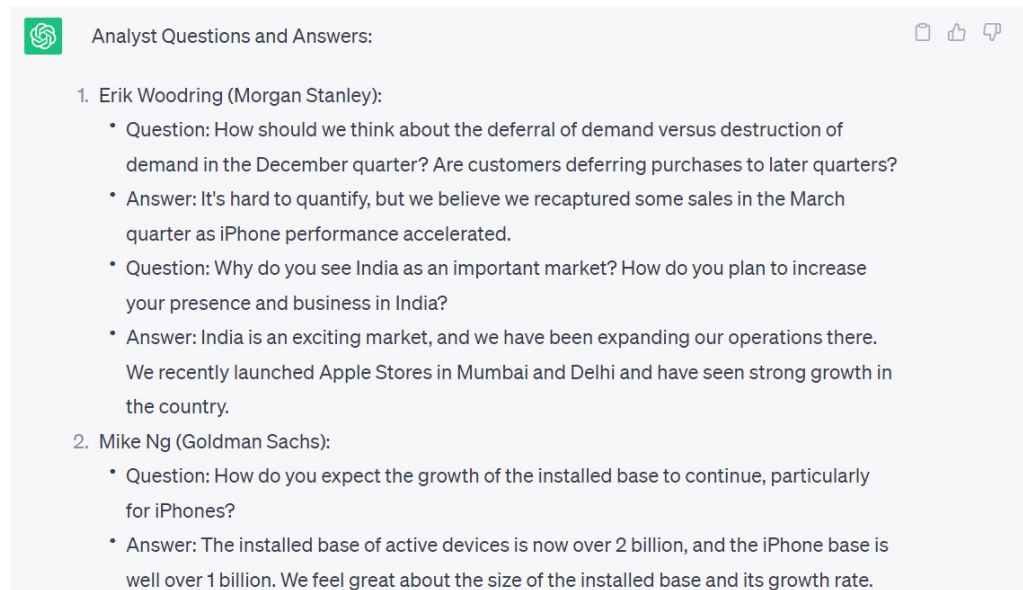
- Objectives:
 - to investigate the **hidden patterns** exhibited by analysts' questions during earnings calls; to measure **sentiment** and level of **forward-looking** content in the questions posed by analysts
- Potential contributions:
 - inform regulators on fairness and transparency in the earnings call Q&A section
 - contribute to the literature in the fields of analysts' behavior and NLP research

Hypotheses

- H1: There exists **bias** in the choice of analysts during earnings calls, influenced by factors such as shareholding and brokerage size.
- H2: The questions selected earlier during the call tend to be more **positive tone** in comparison to subsequent questions.
- H3: The questions selected earlier during the call tend to have a higher degree of **forward-looking content** compared to subsequent questions.

Data and Methodology

- Conference call Q&A transcripts
 - FactSet
 - Seeking Alpha
 - Motley Fool
 - Factiva
- ChatGPT
- FinBERT



The screenshot shows a chat interface with a green OpenAI logo icon. The title is "Analyst Questions and Answers:" and there are icons for copy, like, and reply. The content is a list of two analyst questions and answers:

1. Erik Woodring (Morgan Stanley):

- Question: How should we think about the deferral of demand versus destruction of demand in the December quarter? Are customers deferring purchases to later quarters?
- Answer: It's hard to quantify, but we believe we recaptured some sales in the March quarter as iPhone performance accelerated.
- Question: Why do you see India as an important market? How do you plan to increase your presence and business in India?
- Answer: India is an exciting market, and we have been expanding our operations there. We recently launched Apple Stores in Mumbai and Delhi and have seen strong growth in the country.

2. Mike Ng (Goldman Sachs):

- Question: How do you expect the growth of the installed base to continue, particularly for iPhones?
- Answer: The installed base of active devices is now over 2 billion, and the iPhone base is well over 1 billion. We feel great about the size of the installed base and its growth rate.

Next steps

- develop a stronger theoretical background for each hypothesis
- examine the available datasets and start data collection

Questions and comments



The background of the slide features a large, faint watermark of the Rutgers University seal. The seal is circular with a sunburst in the center and the words "RUTGERS UNIVERSITY" around the perimeter. The word "RUTGERS" is written in a large, red, serif font at the top left of the slide.

RUTGERS

THE STATE UNIVERSITY
OF NEW JERSEY

Do employees know what would happen? Signals for the firm performance derived from the bottom tone

Hanxin Hu, Rutgers

Kelly Duan, Sacred Heart University

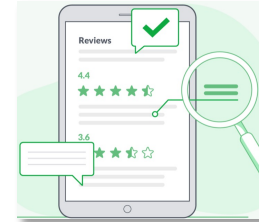
Ting Sun, The College of New Jersey

Motivations and Research Questions

Motivations

- Inside Employee perceptions are indicative of the firm’s prospective financial and non-financial performance.
 - Anecdotal evidence (e.g., Wells Fargo fraudulent account scandal in 2016).
 - Empirical evidence (e.g., Hales et al., 2018; Campbell and Shang, 2021; and Briscoe-Tran, 2022), *where inside views are extracted from a company-reviewing social media platform Glassdoor.*
- Limitations
 - Mostly relying on the homogenous scaled ratings; or using dictionary-based approach to extract textual information.
 - Using the single aggregated-level index to proxy for employee perception
 - The role of “top tone” is often ignored.

Research Questions



Glassdoor textual reviews



Firm financial and non-financial outcome

- How to construct semantically reliable topic features to predict the firm future performance?
- Which aspects of information extracted from Glassdoor reviews are more important?
- Compared to top management tones in the mandatory disclosure, do inside views carry unique and significant information regarding the firm outlook?

Research Design

Step 1: construct topic features

An unsupervised neural network-based topic model (i.e., Contextualized Topic Model from Federico et al, 2021) is developed to classify over 3,200,000 employee review contents into various categories, including career growth opportunities, work-life balance, workplace atmosphere, and compensation benefits.

Step 2: rolling-window forecast

Using a rolling-window approach for the sample period 2010 to 2021, ridge regression models are used to evaluate if employee perceptions contribute incrementally to forecasting firm profitability, net cash flow activities, and non-financial performance. Furthermore, the attribute importance analysis is conducted to determine which topics are strongly linked to firm-level performance

Step 3: Top tones versus bottom tones

We examine whether employee perspectives can complement the forward-looking tones of top management in the MD&A section to augment the accuracy of performance forecasts.

Thanks!

Any comments or suggestions are welcomed and appreciated 😊

hh497@scarletmail.rutgers.edu



The background of the slide features a large, faint, circular seal of Rutgers University. The seal contains the text 'RUTGERS UNIVERSITY' and 'THE STATE UNIVERSITY OF NEW JERSEY' around its perimeter, with a central sunburst design.

RUTGERS

THE STATE UNIVERSITY
OF NEW JERSEY

ESG Report Data Extraction: Generalized Text-Mining Framework

Presented by Steven Katz

Introduction

- ESG reports are long, unstructured, and difficult to interpret/compare
- Considerable array of options for extracting meta-information
 - Sentiment, similarity, part-of-speech prevalence, etc.
- Extracting specific information more difficult
 - Stated goals, reporting framework(s) used, assuring party

Introduction

- Used text mining to answer six questions
 - Want to use experience to generate a widely applicable process
 - Lots of other questions unasked
- Design science project
 - Given a corpus of ESG reports, how to extract information?
 - Deliverable should be a framework that generalizes ESG report text mining
- Target is “researchers who want to use ESG reports but can’t”
 - We don’t claim to propose the only right way
 - Just a sufficiently right way

Goals of ESG Report Text Mining Framework

- What steps must precede analysis?
- What technical pitfalls should be avoided?
- What about ESG reports presents a unique challenge?

What steps must precede analysis?

- Collection/Curation
 - Where to gather data
 - Linked documents
 - Scanning to remove erroneous downloads
- Cleaning/Trimming
 - Strategies and rationale
 - Identifying and removing extraneous pages

What technical pitfalls should be avoided?

- Formatting irregularities
 - Tables
 - Columns
 - Images
 - Encryption
- OCR vs Extraction
 - Pros and cons of each

What about ESG reports presents a unique challenge?

- Lack of standardized terminology
 - Iterative process for building custom bag of words
 - Fuzzy searches
 - Context-sensitive searches
- Limitations of presence/absence as signifier
 - “Absence of evidence is not evidence of absence”
 - Negative language
 - Filer’s incentive to obfuscate

Tentative Framework

- Collection/Cleaning
- Choosing a search method
- Customizing search tools
- Testing/Refinement
- Verification



The background of the slide features a large, faint watermark of the Rutgers University seal. The seal is circular with a sunburst in the center and the text 'RUTGERS THE STATE UNIVERSITY OF NEW JERSEY' around the perimeter.

RUTGERS

THE STATE UNIVERSITY
OF NEW JERSEY

AIS Literature Review Generation

Keywords: *Artificial Intelligence; Deep Learning; Accounting; Automation; Language Model*

Maksym Lazirko

Hanchi Gu

Goals:

- Establish & Utilize Carlab LLM
- Automate the Literature Review Process
- Continuous Auditing Research

Phase I. Experiment

- Involves tinkering with various algorithms/tools
- Needs to allow for (AIS) field specification
- Paper Must have legitimate Citations

Abstract:

As the demand for AI-AI communication increases, the development of an AI internet protocol using natural language has become a hot topic. In this paper, we examine the potential benefits and challenges of such a protocol and explore the existing research in this area. Our analysis suggests that the use of natural language can significantly enhance the efficiency and effectiveness of AI-AI communication. We also highlight the key issues that need to be addressed in order to make such a protocol a reality.

Introduction:

One of the critical challenges in AI development is enabling effective communication between different AI systems. Currently, most AI communication relies on a predefined format or protocol, such as JSON or XML. However, these protocols have several limitations. First, they require some level of technical expertise to use, making it challenging for non-experts to communicate with AI systems. Second, they may not be flexible enough to accommodate the diverse needs of different AI systems. Natural language, on the other hand, is a flexible and intuitive means of communication that can be used by anyone, regardless of technical expertise. In this paper, we explore the potential benefits and challenges of using natural language as an AI internet protocol for AI-AI communication.

Benefits of Using Natural Language:

One of the critical benefits of using natural language as an AI internet protocol is that it can significantly enhance the efficiency and effectiveness of AI communication. By using natural language, AI systems can communicate more easily and intuitively, reducing the need for technical expertise. Moreover, natural language can enable AI systems to communicate in a human-like manner, making it easier for humans to interact with AI systems. Finally, natural language can make it easier to develop complex AI applications that require interaction between multiple AI systems.

Challenges of Using Natural Language:

One of the significant challenges of using natural language as an AI internet protocol

Phase II. Survey

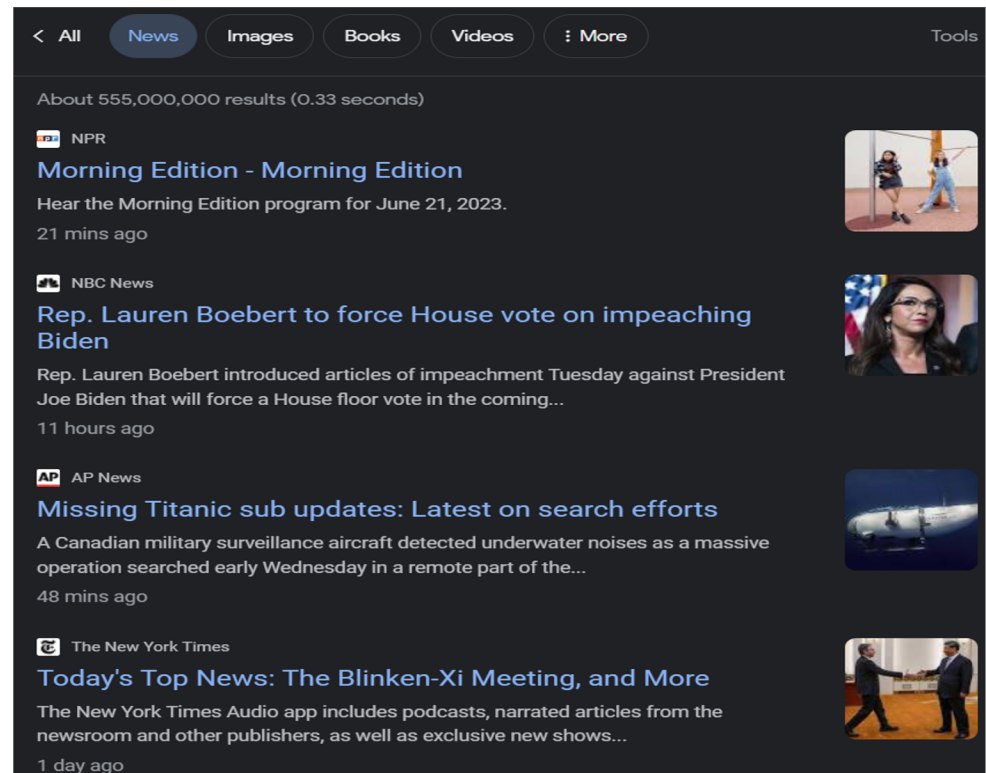
- Recreate sample paper that is outside of training data
- Compare sample paper and its generated mirror
- AIS experts rate quality of generated paper

1. Cho, K. et al., 2014. Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation. arXiv preprint arXiv:1406.1078.
2. Dai, A. M., Oliva, J. & Le, Q. V., 2015. Semi-supervised sequence learning. In Advances in neural information processing systems, pp. 3079–3087.

-
3. Mikolov, T. et al., 2013. Efficient estimation of word representations in vector space. arXiv preprint arXiv:1301.3781.
 4. Pennington, J., Socher, R. & Manning, C. D., 2014. Glove: Global vectors for word representation. In Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP), pp. 1532–1543.
 5. Vaswani, A. et al., 2017. Attention is all you need. In Advances in neural information processing systems, pp. 5998–6008.

Phase III. Adjustments & Future Research

- Evaluate survey and address possible fixes/limitations
- Web Scraping news to determine relevant research topics
- Use current events as a seed for algorithm



Thank
You



CONTINUOUS MONITORING SYSTEM WITH INTERACTIVE VISUALIZATION: CASE STUDY ON GOVERNMENT PAYROLL DATABASE

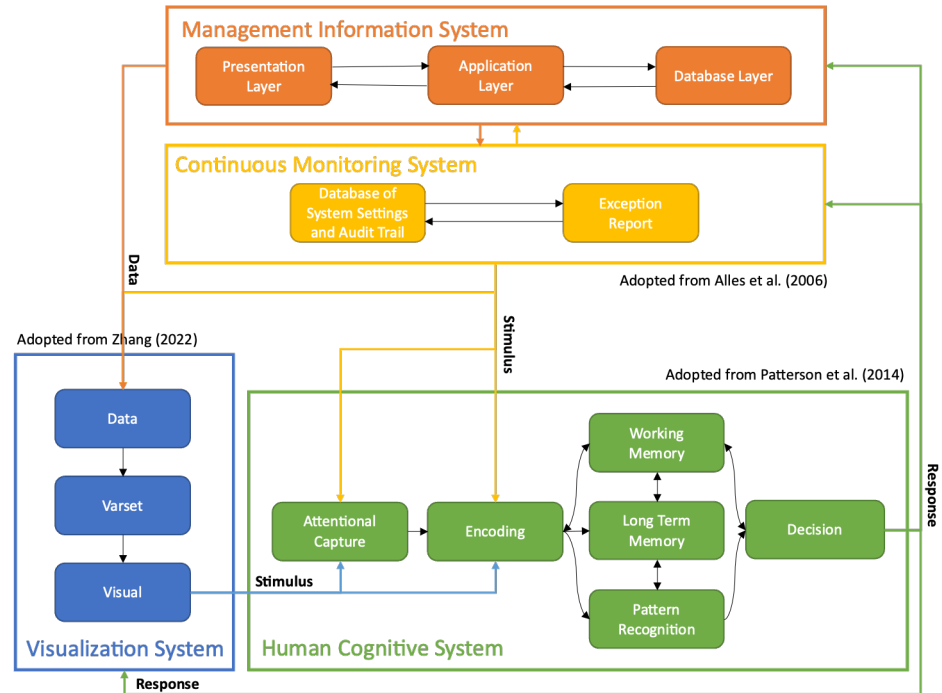
Heejae (Erica) Lee
Ph.D. Candidate

Introduction

- Several studies showed that visual analytics can facilitate decision-making by improving cognitive effort productivity.
- Integrating interactive data visualization into the CM system can improve decision-making by allowing users to communicate more efficiently with the reported results.
- Examine the role of visualization in interactions between human cognitive and continuous monitoring systems.

Model

- Management Information System & Continuous Monitoring System (Adopted from Alles et al. 2006)
- Visualization System (Adopted from Zhang 2022)
- Human Cognitive System (Adopted from Patterson et al. 2014)



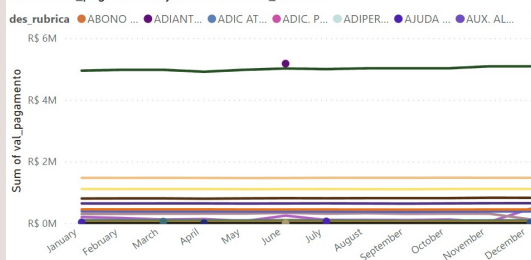
Case Study

- Work with internal audit team to develop visual analytic dashboard
- Data were extracted from the payroll database of the state government of Santa Caterina in Brazil.
- Payroll data reported in 2021 were used to develop a prototype of the continuous monitoring dashboard.
- Data includes 26,107,538 records of 169,082 current and retired employees from 46 departments under the state government.

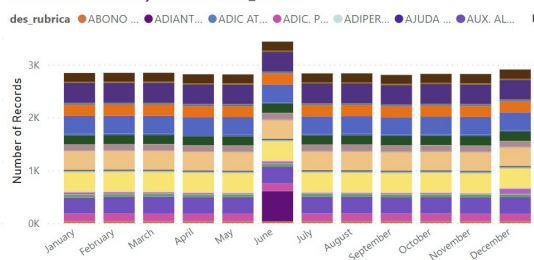
Current CA system

- The CA system creates reports for payment where the change is greater than cutoff every month.
- Have two business days before the payment is processed.
- The initial cutoff was set to monthly change greater than 3,000 Brazilian Real (\$625)
- The cutoff was updated to 30,000 Brazilian Real (\$6,250) due to volume of exceptions

Sum of val_pagamento by Month and des_rubrica



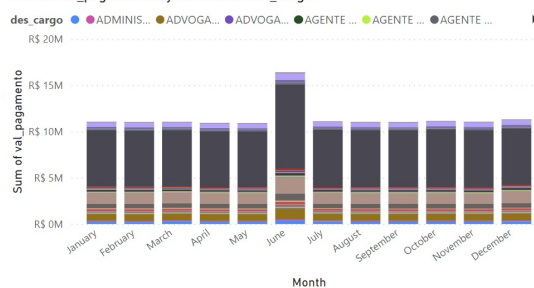
Number of Records by Month and des_rubrica



- des_rubrica
- ABONO DE PERMANENCIA
 - ADIANT 13 SALARIO
 - ADIC. POS-GRADUACAO
 - ADIPERMANENCIA 5%
 - AJUDA DE CUSTOS
 - AUX. ALIMENTACAO PROV
 - COMPLEMENTO PISO DO MAGISTERIO
 - DIF. GRAT. FERIAS
 - FERIAS INDENIZADAS
 - FUNCAO GRATIF FG

- Month
- ADIANT 13 SALARIO
 - AJUDA DE CUSTOS
 - GRAT FERIAS CONST
 - INDENIZACAO DE USO DE VEIC...

Sum of val_pagamento by Month and des_cargo



Design of Visual Analytics Dashboard

- If the monthly *Change* of a certain payment type is greater than 3,000 Brazilian Real, it will be listed in the slicer.

Future Plan

- Complete prototype and share it with internal audit team
- Receive feedback from internal audit team and revise the dashboard



The background of the slide features a large, light gray watermark of the Rutgers University seal. The seal is circular with a sunburst in the center and the text 'RUTGERS THE STATE UNIVERSITY OF NEW JERSEY' around the perimeter.

RUTGERS

THE STATE UNIVERSITY
OF NEW JERSEY

Enhancing Payroll Fraud Detection for Governments: Continuous Auditing with Deep Learning based Outlier Detection

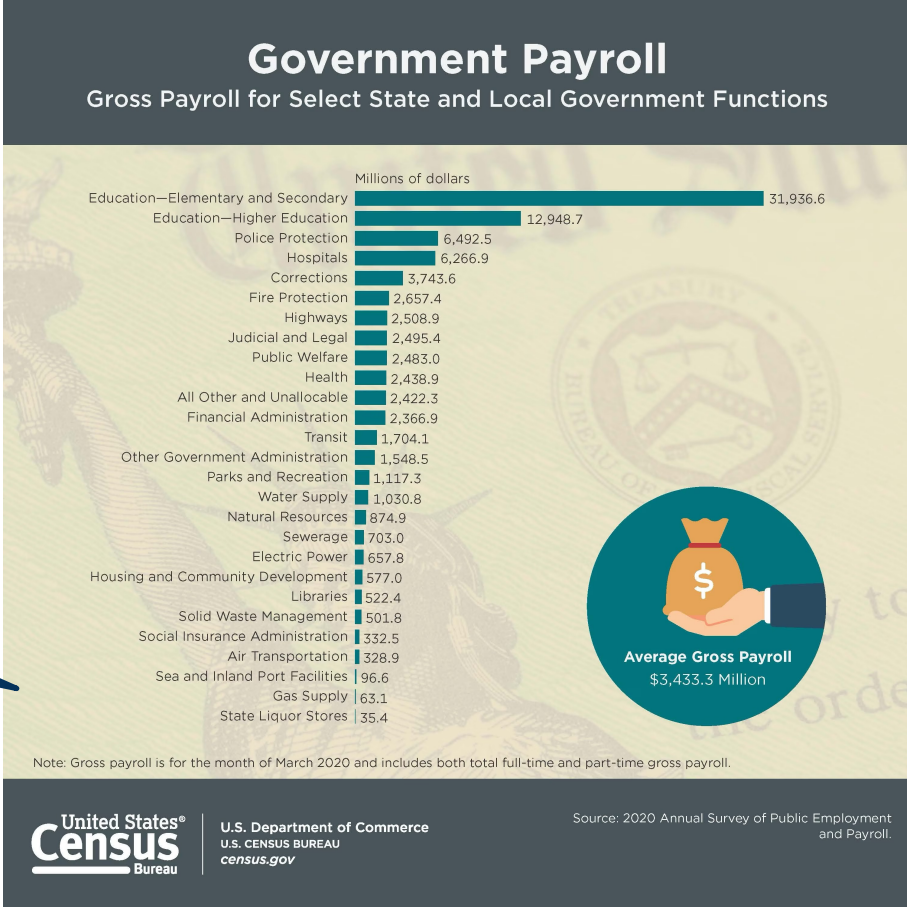
Huaxia Li

Summer Research 2023

Motivation

Government programs are financed by payments from taxpayers, and as a result, it is essential to ensure that their funds are expended appropriately (Mead, 2008, Rozario and Issa, 2020)

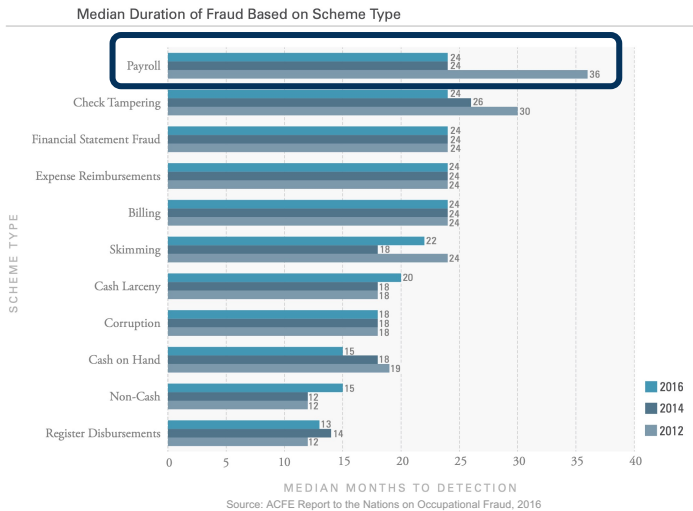
89 billion for March 2020



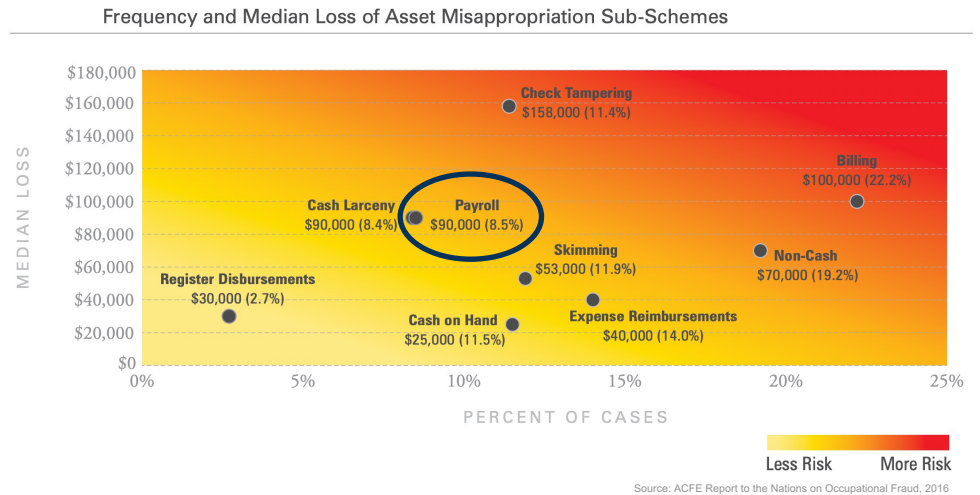
Source: Government Payroll <https://www.census.gov/library/visualizations/2021/icon/government-payroll.html>
 Mead, D. M. (2008). SEA performance reporting. *The CPA Journal*, 78(1), 4.
 Rozario, A. M., & Issa, H. (2020). Risk-based data analytics in the government sector: A case study for a US county. *Government Information Quarterly*, 37(2), 101457.

Motivation

- Payroll fraud is **prevalent** in the non-profit sector (Wells, 2005) and is among the **top five** major types of fraudulent disbursements (Greenlee et al. 2007).

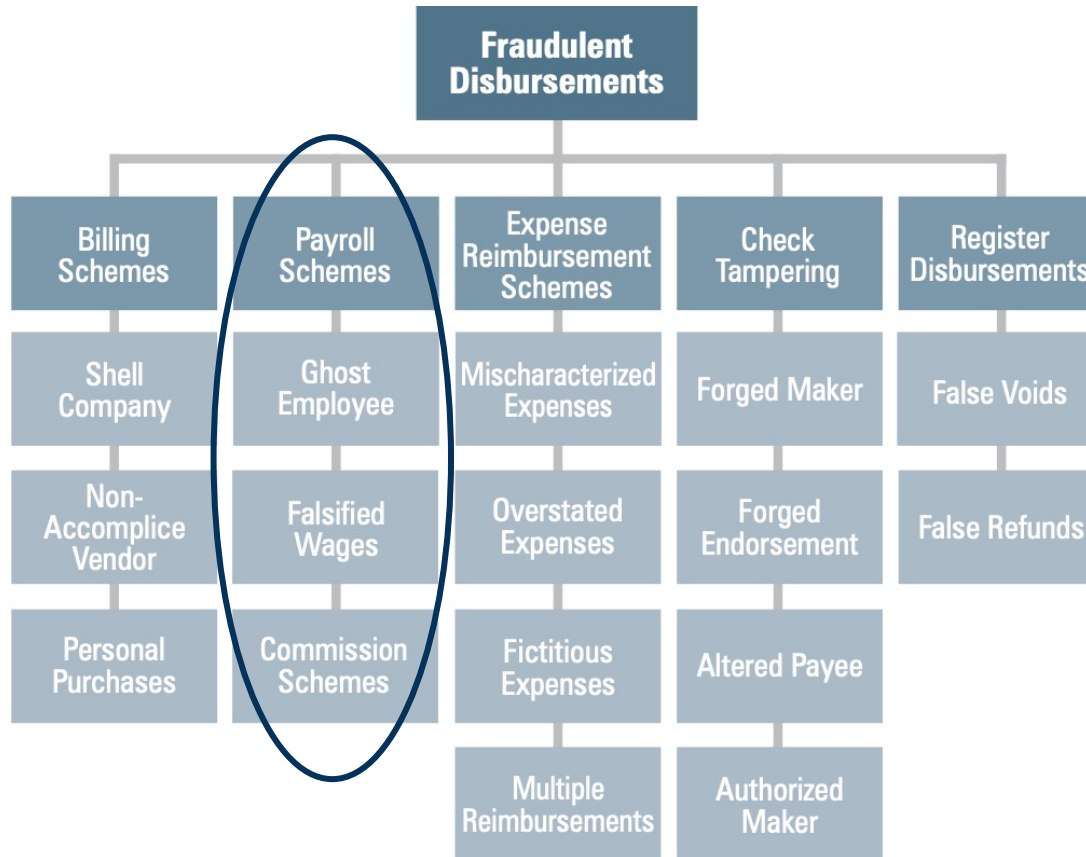


etc



Source:
 Wells, J.T. (2005). Principles of Fraud Examination. John Wiley, New York, NY, N.Y.
 Greenlee, J., Fischer, M., Gordon, T., and Keating, E. (2007). "An investigation of fraud in nonprofit organizations: occurrences and deterrents", Nonprofit and Voluntary Sector Quarterly, Vol. 36 No. 4, pp. 676-694.
 REPORT TO THE NATIONS ON OCCUPATIONAL FRAUD AND ABUSE 2016 GLOBAL FRAUD STUDY (<https://www.acfe.com/~/media/files/acfe/pdfs/2016-report-to-the-nations.ashx>)

Motivation



Introduction

- Internal auditors are facing the **challenges** of working with large, instantly accessible data generated continuously and automatically within the government's payroll systems¹.
- In Brazil, governmental organizations can **exchange information** with other governmental organizations and use these as external databases².
- In Brazil, public servants are usually **stable** in their jobs².
 - For instance, after three years in the job, it is very difficult for a public servant to be fired, with very few exceptions (involvement in corruption, for example).

Source:

1. Eulerich, M., & Kalinichenko, A. (2018). The current state and future directions of continuous auditing research: An analysis of the existing literature. *Journal of Information Systems*, 32(3), 31-51.

2. de Freitas, M. M., Codesso, M., & Augusto, A. L. R. (2020). Implementation of continuous audit on the Brazilian navy payroll. *Journal of Emerging Technologies in Accounting*, 17(2), 157-171.

- Continuous audit (CA) with advanced technologies is suitable for identifying hidden fraud in these large, complex, and

Introduction

- Outlier detection (OD) methods are **promising** in detecting anomalies from individual transactions or transaction sequences (Zhang et al., 2022).
- OD techniques, such as cluster analysis, can be integrated into a schema of continuous auditing and continuous monitoring benefit from their **unsupervised** features (Thiprungsri and Vasarhelyi, 2011).
- This study will investigate how to enhance payroll fraud detection in government sectors through continuous auditing with OD techniques. **Especially**, how the advanced method outperforms traditional methods.

Source:
Zhang, C., Atasoy, S., & Vasarhelyi, M. A. (2022). Continuous auditing with machine learning and interaction visualization: An application to a healthcare payroll process. *International Journal of Accounting Information Systems*, 46, 100577.
Thiprungsri, S., & Vasarhelyi, M. A. (2011). Continuous auditing: An application to a healthcare payroll process. *International Journal of Accounting Information Systems*, 46, 100577.

Data

- Payroll data from one of the states in Brazil
 - State Population: 7,338,473¹ (11th of the 26 states²)
 - Jan. 2021 – Nov. 2021
 - 5,863,855 payment records
 - 76 features (individual level, job level, payment level)
 - 140,880 servants
 - 43 branches
 - 406 positions
 - ...

Source:

1. <http://www.citypopulation.de/en/brazil/cities/santacatarina/>

2. [https://en.wikipedia.org/wiki/Santa_Catarina_\(state\)#:~:text=Santa%20Catarina%20\(Portuguese%20pronunciation%3A%20%5B.and%20the%2011th%20most%20populous.](https://en.wikipedia.org/wiki/Santa_Catarina_(state)#:~:text=Santa%20Catarina%20(Portuguese%20pronunciation%3A%20%5B.and%20the%2011th%20most%20populous.)

Research questions

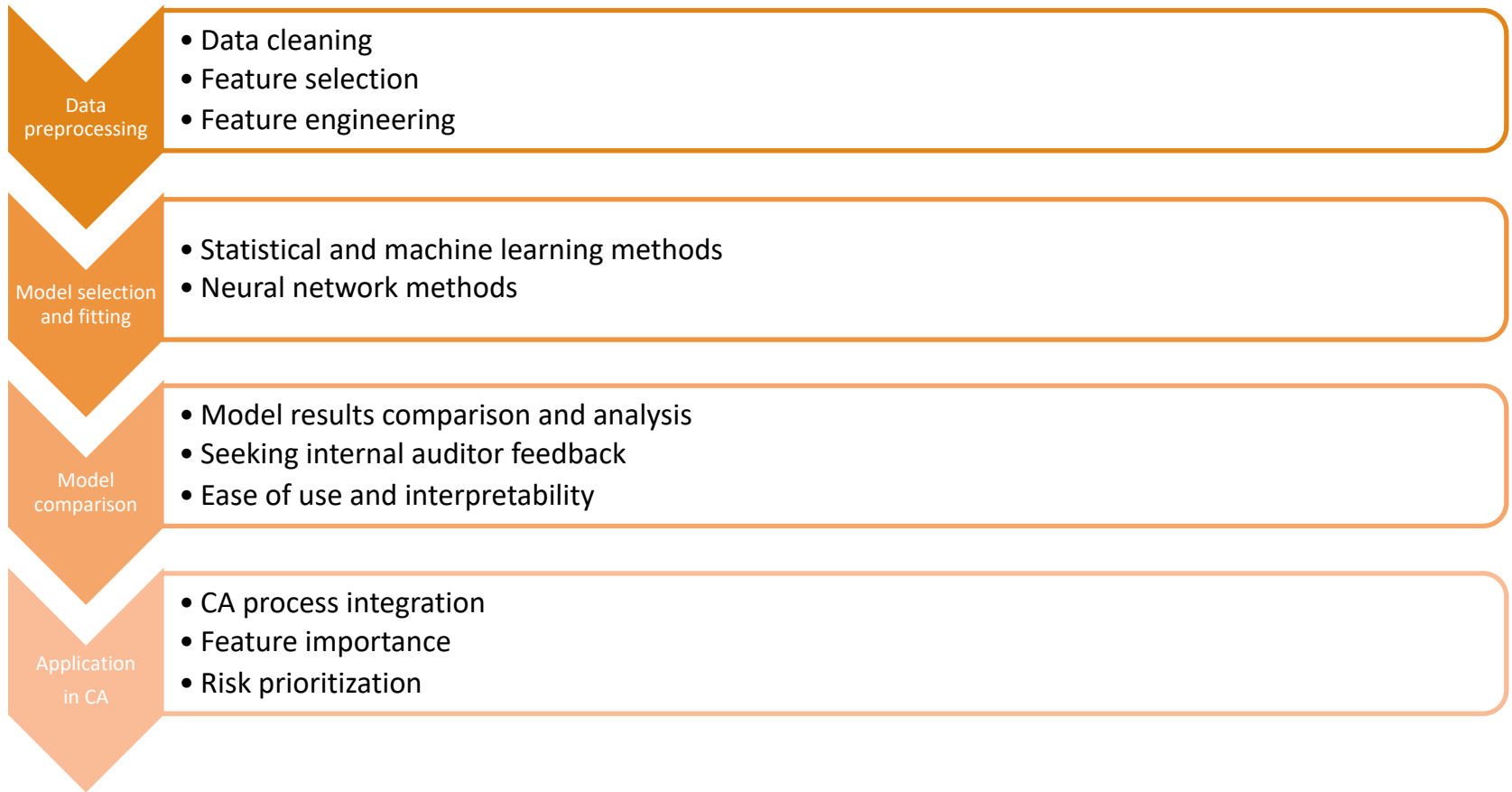
1) What are the characteristics of fraudulent activities in governmental payroll systems?

2) How can OD be applied to identify potential fraud in governmental payroll systems?

3) Whether the advanced OD methods (e.g., deep learning) could outperform traditional methods?

4) What are the benefits and challenges of applying CA and OD to governmental payroll systems?

Research design



Next steps

- **Model selection and fitting**
 - Select models from different categories
 - Fitting the data with the models, summarize and visualize the results.
- **Model comparison**
 - Comparing the results from advanced model result with the those from the benchmark model
 - Analyzing the ease of use and cost through consulting the state internal auditors
- **Application in CA**
 - Risk prioritization
 - Benefits and limitations analysis

THANK YOU



The background of the slide features a large, faint watermark of the Rutgers University seal, which is a circular emblem with a sunburst design and the text 'THE STATE UNIVERSITY OF NEW JERSEY' around the perimeter.

RUTGERS

THE STATE UNIVERSITY
OF NEW JERSEY

ESG violation and insider trading

Nichole Li

Rutgers , The State University of New Jersey

Meehyun Kim

Rutgers, The State University of New Jersey

Min Cao

University of Northern Colorado

Miklos Varsarhelyi

Rutgers , The State University of New Jersey

Motivation

- The general public and investment community rely heavily on ESG scores/ratings provided by third-party commercial data vendors to identify ESG-friendly firms
- Growing evidence suggests that due to information asymmetry and the lack of assurance, having a decent ESG reputation does not translate into socially responsible practices.

Motivation

- Stock market participants want to know if insider trading is widespread because it affects investors' willingness to trade, and consequently affects the liquidity of the stock.
- Policy makers and regulators concerned about the effectiveness of existing insider trading regulations.

Research questions

- 1. Whether a firm's own social and environmental violations lead to lower ESG ratings.
- 2. Whether insiders sell ESG violation firms' shares before the news of the violation is announced to the general public.

Data

- Violation tracker
- CRSP
- Refinitiv

Data

- 2020/01/01-2022/03/22

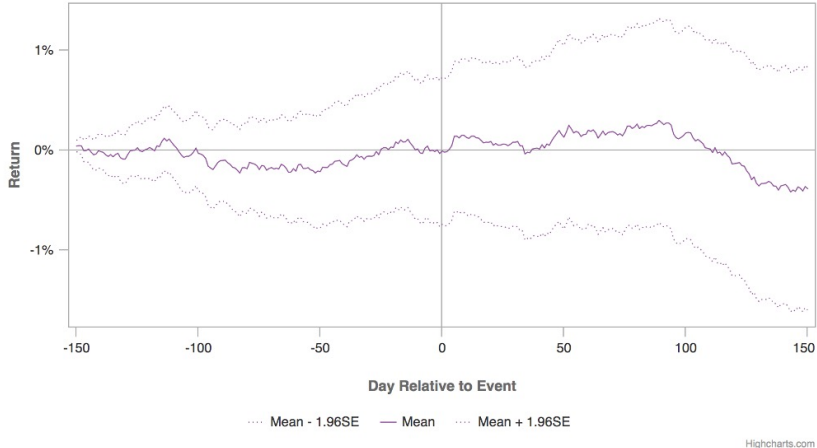
Number of records in total	512,627
Number of records have ticker	67,633
Number of records have ticker and Total asset	62,824
Number of records after aggregation	46,363

Event study

- Penalty amount scaled by the total assets
- 10 percentile

Cumulative Abnormal Return: Mean & 95% Confidence Limits

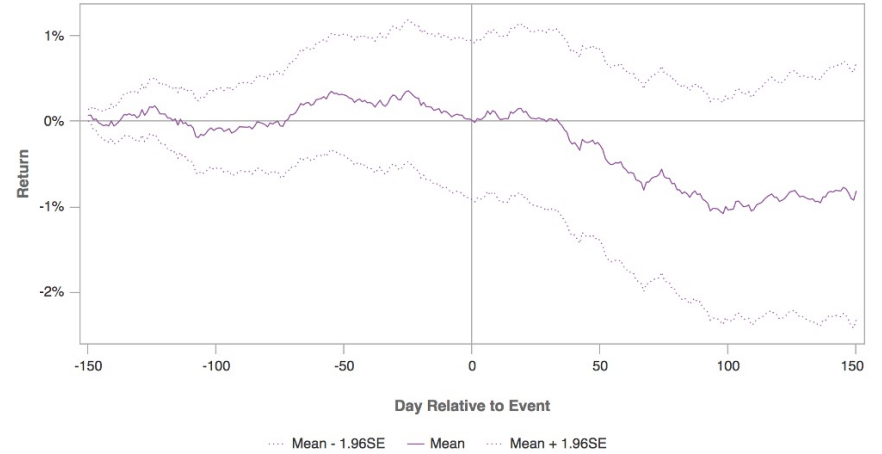
There are 3631 events in total with non-missing returns.



Highcharts.com

Cumulative Abnormal Return: Mean & 95% Confidence Limits

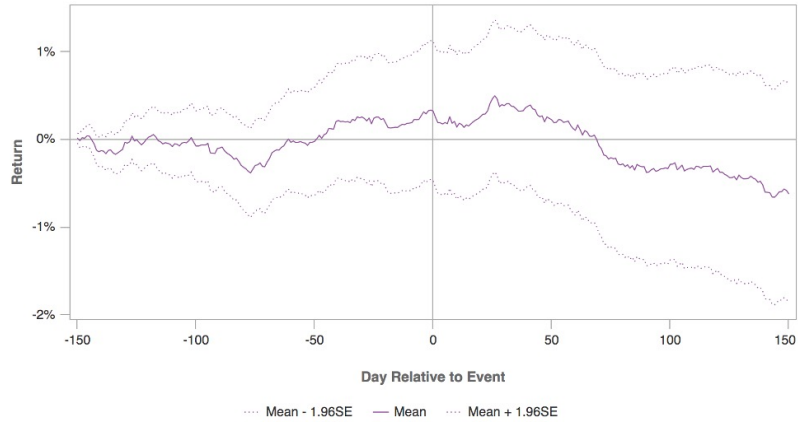
There are 3955 events in total with non-missing returns.



Highcharts.com

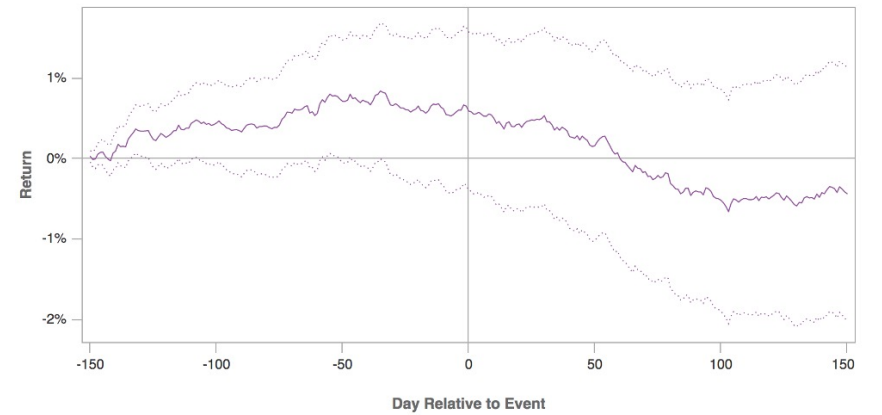
Cumulative Abnormal Return: Mean & 95% Confidence Limits

There are 3998 events in total with non-missing returns.



Cumulative Abnormal Return: Mean & 95% Confidence Limits

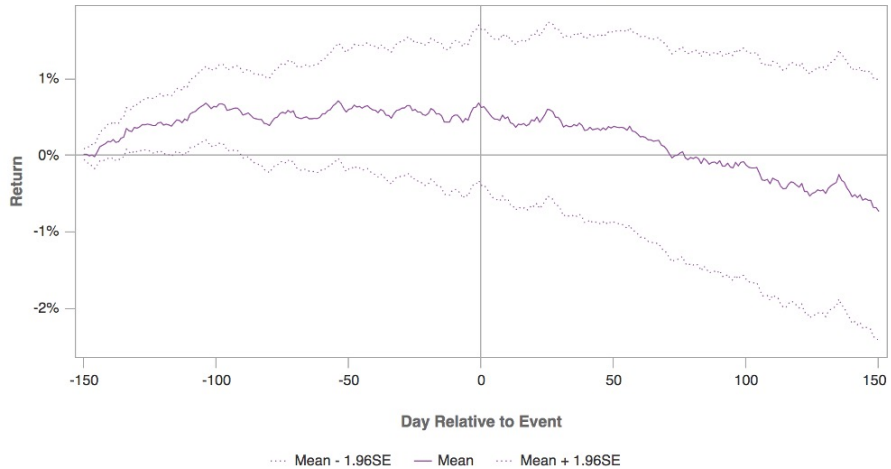
There are 4007 events in total with non-missing returns.



Day Relative to Event

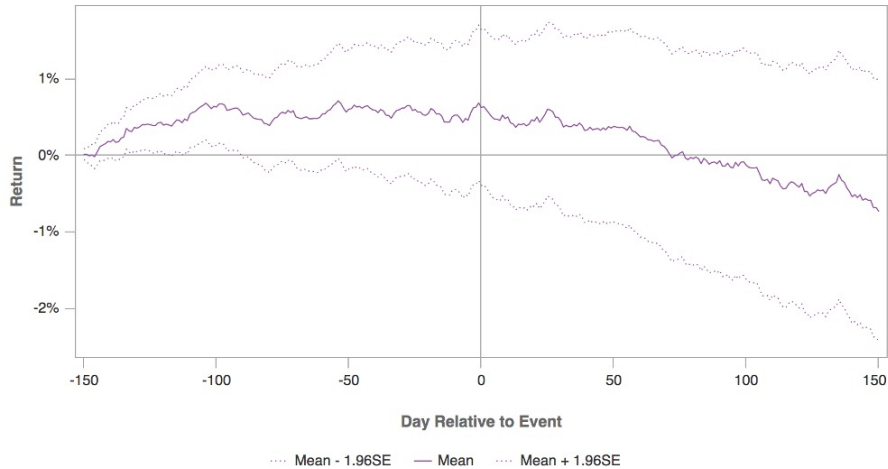
Cumulative Abnormal Return: Mean & 95% Confidence Limits

There are 3981 events in total with non-missing returns.



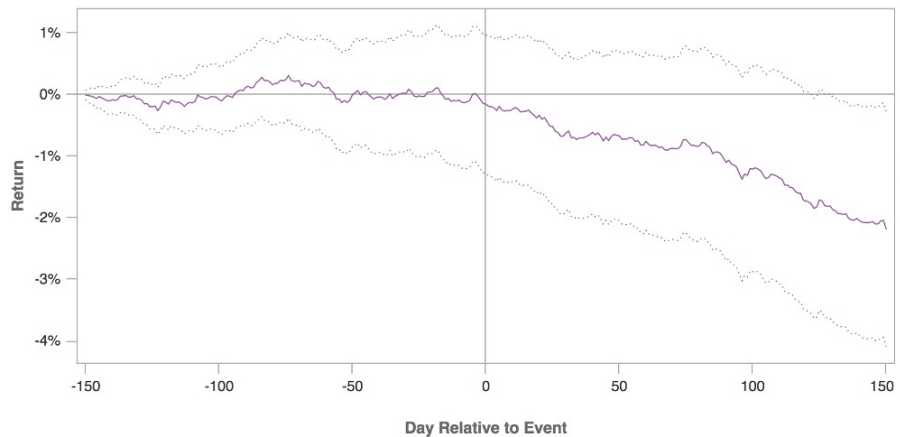
Cumulative Abnormal Return: Mean & 95% Confidence Limits

There are 3981 events in total with non-missing returns.



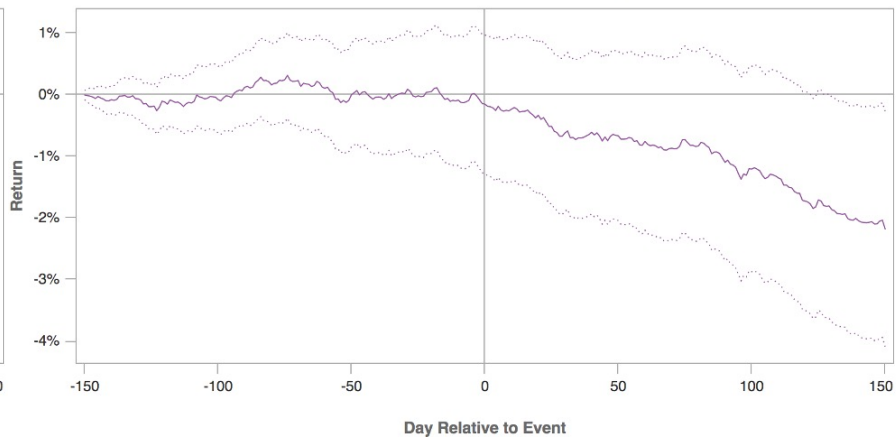
Cumulative Abnormal Return: Mean & 95% Confidence Limits

There are 3804 events in total with non-missing returns.



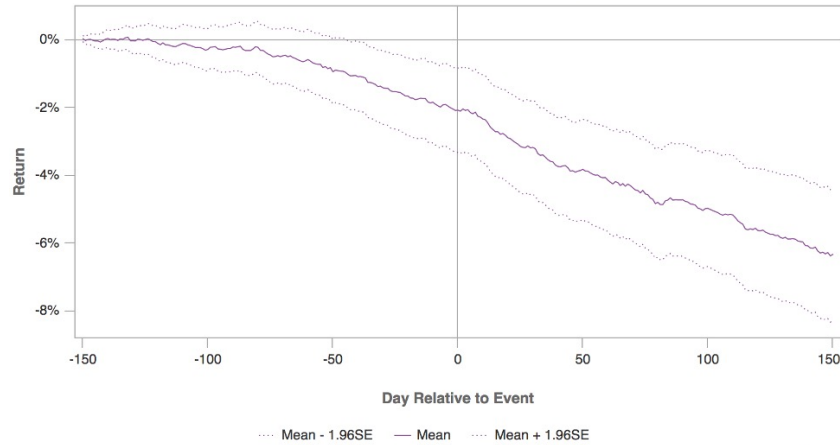
Cumulative Abnormal Return: Mean & 95% Confidence Limits

There are 3804 events in total with non-missing returns.



Cumulative Abnormal Return: Mean & 95% Confidence Limits

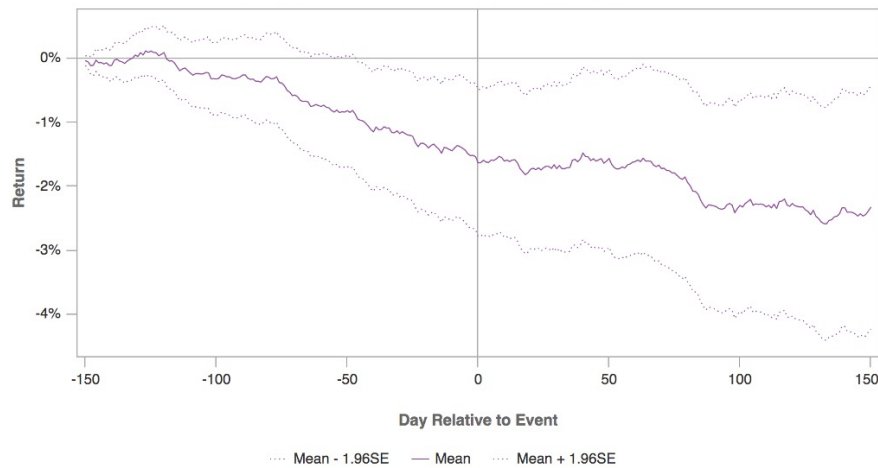
There are 3762 events in total with non-missing returns.



Highcharts.com

Cumulative Abnormal Return: Mean & 95% Confidence Limits

There are 3788 events in total with non-missing returns.



The background of the slide features a large, light gray watermark of the Rutgers University seal. The seal is circular and contains a sunburst in the center, with the words "RUTGERS UNIVERSITY" and "THE STATE UNIVERSITY OF NEW JERSEY" around the perimeter.

RUTGERS

THE STATE UNIVERSITY
OF NEW JERSEY

Summer Project Progress

Shaoyu Liu

Municipality Government Reporting Automation

- Government Finance Officers Association (GFOA)
 - A nonprofit organization that represents public finance officials throughout the United States and Canada
 - 20,000 members of federal, state/provincial, and local finance officials
 - involved in planning, financing, and implementing thousands of governmental operations
- Robotic Process Automation (RPA) and Public Reporting
 - Standardized and repetitive tasks
 - Cannot deal with uncommon and unique cases

Municipality Government Reporting Automation

- Objectives
 - Looking for commonalities among different reporting schemes
 - Create an RPA routine library
 - Indicate what public reporting tasks are feasible for automation
 - Create a protocol
 - Help and instruct users to figure out what tasks and how to automate

Municipality Government Reporting Automation

- Current Stage
 - Very early initiate stage
 - Exploring the feasibility and scope of this project
- Define criteria for automation tasks
 1. Repetitive
 2. Resource intensive (High potential saving)
 3. Substantial stakes (Enough stake for people to be interested in)
 4. Definable (Documents the steps)
 5. Rule-based
 6. Machine-readable inputs

Payroll / Time &
Attendance

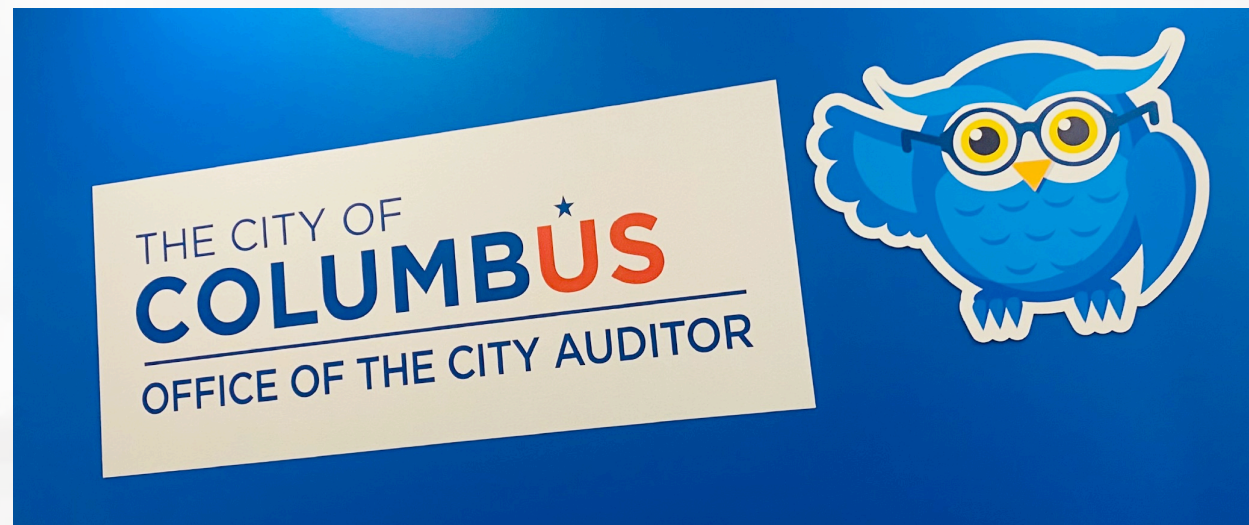
ERP / Microsoft
D365

Revenue System

Transparency
Portal

Banking &
Treasury

City of Columbus: Digital Transformation



Municipality Government Reporting Automation

- Columbus, OH – early mover
 - Routine tasks (e.g. preparing and sending letters)
 - Organizing information (e.g. PPP loan distribution)
 - 24/7 assistance to customers (e.g. chatbots, AI-helpers)
 - Algorithmic processes (e.g. audit determination)
 - Trend analyses (e.g. revenue projections)
 - Operations (e.g. smart safes, tax form uses)
- CTO: “I think of it as like having another person on the team”

Municipality Government Reporting Automation

- Future Plan
 - Contact Columbus Automation Team to have a show & tell
 - Understand what they have done and what we can learn
 - Whether their experience is replicable
 - Figure out what we can build up on their work
 - Subject of automation
 - Not only financial reporting, expanding to operation and auditing
 - Scope of automation
 - Baseline consistency of tasks for different municipalities also using Microsoft Dynamic 365
 - Extended to other ERP systems like Oracle and SAP



The background of the slide features a large, faint watermark of the Rutgers University seal. The seal is circular with a sunburst in the center and the text 'RUTGERS THE STATE UNIVERSITY OF NEW JERSEY' around the perimeter.

RUTGERS

THE STATE UNIVERSITY
OF NEW JERSEY

Predicting revenues using customer reviews

Presenter: Talha

- Source of reviews: Amazon
- Data elements: Product id, review id, parent category, review title, review
- Data made available by: Stanford University

Plan

- Model topics using LDA.
- Assign sentiments to review by using star ratings as a guide
- Less than or equal to 2 stars = negative review
- Greater than or equal to 3 stars = positive review
- Correlate topics and sentiments with sales revenue

Plan

- Sales revenue will be proxied using Amazon sales rank.
- Higher sales rank = higher sales
- Prior study shows strong correlation between sales rank and sales revenue.
- Other studies have used similar proxies for sales revenue. For example, one study uses the number of customer reviews to proxy for sales revenue.

Plan

- I am trying to do a large scale study with as many categories and reviews as I can.
- Possibly over a million textual reviews.

Current studies:

- E-Word of mouth & Box office revenue
- Volume of reviews & sales on Amazon (Cui et al 2012)
- Average rating and sales on Amazon (Cui et al 2012)
- Some studies just look at certain categories

Software end

- Loading csv files line by line into python memory
 - Processing each line before moving on to the next
- Threading in python – for parallel computing:
 - I will create multiple instances of the code to run simultaneously
- I will use SQL database for storing work in progress directly to hard disk.
 - Work in progress includes processed text.
- Small chunksize for LDA training will help keep memory usage down.

Methodology

- Studies have used time series analysis.
- I will use a fixed effect analysis as I am looking for long term effects.
- Control variables (borrowed from previous studies):
 - Price of product
 - Product type
 - Product subcategory

Thankyou



The background of the slide features a large, faint watermark of the Rutgers University seal. The seal is circular with a sunburst in the center and the words "RUTGERS UNIVERSITY" around the perimeter.

RUTGERS

THE STATE UNIVERSITY
OF NEW JERSEY

Continuous Auditing in the Age of AI: An Exploration of OpenAI's GPT Model Applications

Jiaqi Sun Kevin Moffitt

Introduction

- To develop a conceptual framework for continuous auditing using OpenAI's GPT model,
- To understand the potential of AI for real-time risk analysis and anomaly detection in transactional data,
- To examine the challenges and solutions in implementing AI in an auditing context.

Research Overview

- Data Collection and Preparation: Gathering relevant transactional data and preprocessing it to suit our model's needs.
- Fine-tuning the GPT model: Adapting the model to understand and analyze our auditing data effectively.
- Evaluating Model Performance: Rigorously testing our model's accuracy and reliability in risk prediction.
- Implementation: Developing a system that integrates this model into existing auditing systems, and can provide real-time feedback to auditors.

Progress Made

- Week 1 & 2:

Studied fundamentals of continuous auditing and AI.

Explored the specifics of OpenAI's GPT model.

- Week 3 & 4:

Identified challenges in traditional auditing.

Explored the potential role of AI in mitigating these challenges.

- Week 5:

Started investigating types of auditing data required.

Developed data collection methods.

Began working on data preprocessing techniques.

Findings & Insights

- Identified growing interest in AI's potential in auditing
- GPT model seems well-suited for auditing tasks due to its language processing capabilities.
- GPT model could potentially understand transactional data, assess risk levels, and generate real-time reports.
- Major challenges include ensuring interpretability of the GPT model's outputs and the quality of training data.

Challenges & Next Steps

- Challenge 1: Complexity of understanding and applying the GPT model in auditing.
- Solution: Thorough study of the model and consultation with experts.
- Challenge 2: Identifying and acquiring suitable datasets for the study.
- Solution: Reached out to various organizations for data and simulated our own data where necessary.

Next Steps

- Deep dive into risk assessment and explore how AI can facilitate it.
- Fine-tune the GPT model specifically for auditing tasks.
- Evaluate the performance of the AI model in auditing,



The background of the slide features a large, faint watermark of the Rutgers University seal. The seal is circular with a sunburst design in the center and the words "RUTGERS UNIVERSITY" and "THE STATE UNIVERSITY OF NEW JERSEY" around the perimeter.

RUTGERS

THE STATE UNIVERSITY
OF NEW JERSEY

**The Relationship Between the Proportion of Real Outliers and the
Average Attenuation Rate of Outlier Scores: An Approach of
Algorithm Selection in Unsupervised Outlier Detection without
Ground Truth**

Kathy Wei
Alexander Kogan

Introduction

Unsupervised outlier detection

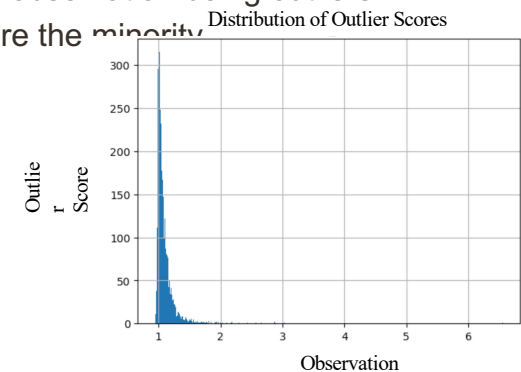
- The essence of unsupervised outlier detection is a process of normality construction and deviation computation achieved without training.
- The normality of a dataset can vary when constructed by different detection methods, which explains why different observations are flagged as outliers.

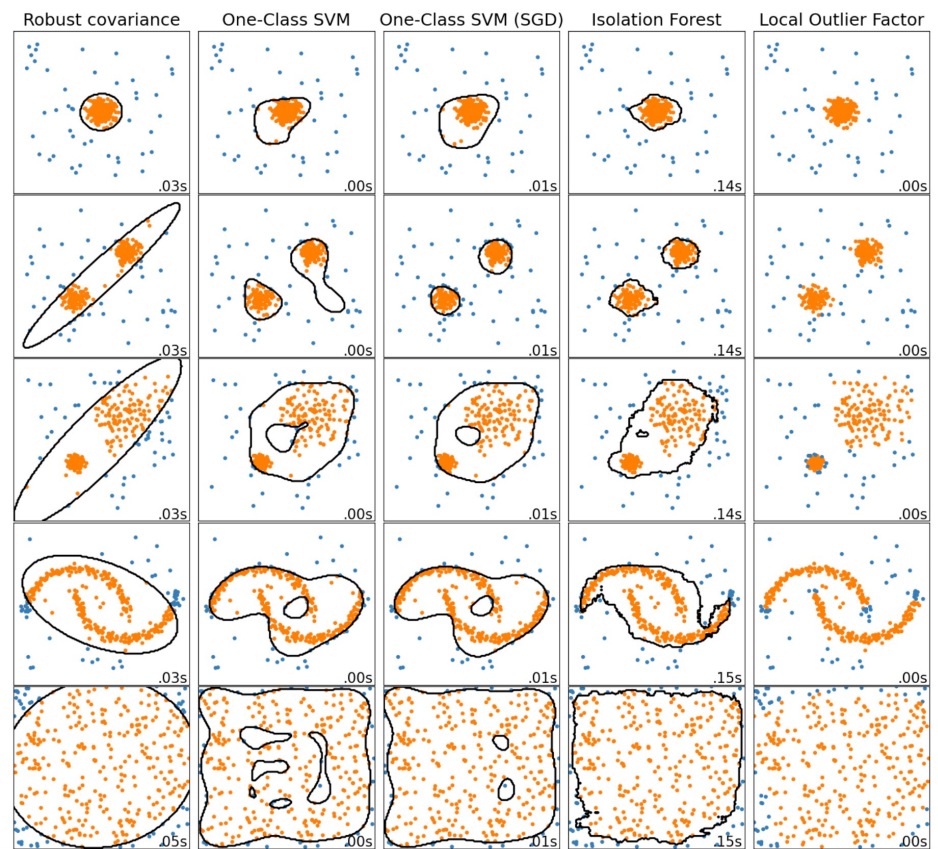
Fitness of a method to a dataset is difficult to evaluate

- Since the ground truth is not known, it is not possible to evaluate how close the constructed normality is to the real one.

The only output: outlier scores

- As the only output, outlier scores indicate the relative probability of an observation being outliers.
- Usually, the distribution of outlier scores is skewed because outliers are the minority.





- Inliers
- Outliers

Dataset 1

Dataset 2

Dataset 3

Dataset 4

Dataset 5

The proportion of outliers for the 5 datasets is 15%.

Decision boundaries between inliers and outliers are displayed in black except for Local Outlier Factor (LOF) as it has no predict method to be applied on new data when it is used for outlier detection.

Image Source: https://scikit-learn.org/stable/auto_examples/miscellaneous/plot_anomaly_comparison.html#sphx-glr-auto-examples-miscellaneous-plot-anomaly-comparison-py

Hypothesis Development

Outlier or inlier: an unobservable score cutoff x_c

- Assuming an unobservable score cutoff x_c exists:
 - A score smaller than x_c indicates an observation is an inlier.
 - More observations that have a score higher than x_c mean more outliers.
- To better illustrate the concept, the outlier scores are normalized to range from 0 to 1 and the unobservable score cutoff after normalization is x'_c .
 - The decrease of outlier scores from 1 to 0 is called attenuation.
 - The decrease rate of outlier scores from 1 to x'_c is called the average attenuation rate of outlier scores.

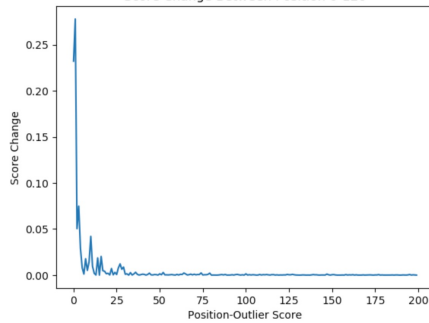
Average attenuation rate of outlier scores vs. Proportion of real outliers

- Hypothesis: The algorithm that has the most observations (and the lowest average attenuation rate) before x'_c fits the dataset best.
- However, x'_c is not known. Hence, the hypothesis is manipulated to make it testable:

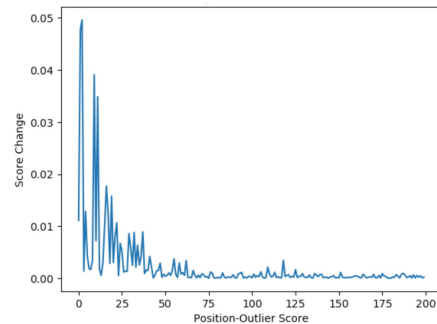
Assume the average attenuation rate of a sample with no outliers is x_0 . As the proportion of outliers in the sample increases, the average attenuation rate will first increase and then decrease.

Hypothesis Development

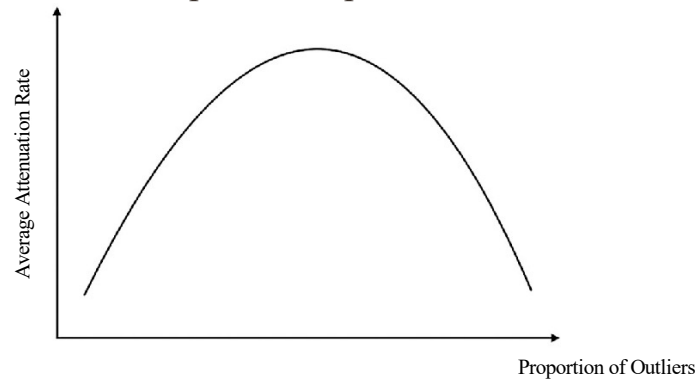
Score Change – MCD (Top 200)



Score Change – LOF (Top 200)



- The average attenuation rate of the sample is equal to the maximum score in the sample minus the minimum score, divided by the sample size.
- According to the manipulated hypothesis, the relationship between the average attenuation rate of the sample and the proportion of outliers in the sample can be presented as an inverted U-shape curve.



Hypothesis Testing

- The manipulated hypothesis is used to develop an algorithm selection tool.
- The tool is applied to select the best algorithm for various datasets.
- If the tool can select the algorithm that identifies the most (or close to to the most) outliers, then the hypothesis is true.

THANK YOU!



The background of the slide features a large, faint watermark of the Rutgers University seal. The seal is circular with a sunburst in the center and the words "RUTGERS THE STATE UNIVERSITY OF NEW JERSEY" around the perimeter.

RUTGERS

THE STATE UNIVERSITY
OF NEW JERSEY

Role-play game (RPG) with an integrated story for accounting education

Fangbing Xiong

Eva Blondeel

Overview

Game-based learning (GBL)

Previous literature has shown that game-based learning (GBL) in accounting education would improve students’:

- Increases their motivation and attitude and perceived learning (Silva & Rodrigues et al., 2021).
- Provides the students with immediate feedback (Kapp, 2012a, 2012b),
- Problem-solving skill improvement (San Cristóbal, 2015; Sung et al., 2015)
- Communication skill improvement (Reinders & Wattana, 2014).
- Helps the students to be the "real agents" of their learning process. (Landers & Armstrong, 2017)

Role-play game (RPG)

Role-play game (RPG) is an application of GBL. RPGs can create stories for students to participate in and allow students to access knowledge in an **integrated way**. (Grande-de-Prado et al., 2020)

Overview

Role-play game (RPG) (Continued)

- Meanwhile, entry-level accounting education also emphasizes **an integrated story**, as we teach the journal entry recording in accounting cycles (e.g., revenue cycle, payable cycle)
- And require students to **make a summary** and adjustments of the journal entry (e.g. prepare a cash flow sheet/ balance sheet at the end of years)
- .RPG can be a good way for accounting education, as well as in the subjects it has been proven beneficial (e.g., Language, history), incorporating with RPGs education tools e.g., Classcraft. (Garcia & Ramon et al., 2010; Lu & Chang, 2016)

Contribution

- The previous common accounting gamifications are Monopoly, Jeopardy, Bingo, Puzzles, Twenty-questions, and some simulations. (Moncada & Moncada, 2014)
- There are few accountings gamification is using integrated stories as RPG.
- As introduction level of financial accounting emphasizes the education target of learning accounting cycle, the research will test **whether the student would benefit from learning accounting in a digital integrated story** by interacting with an education game.

Design

Game Design

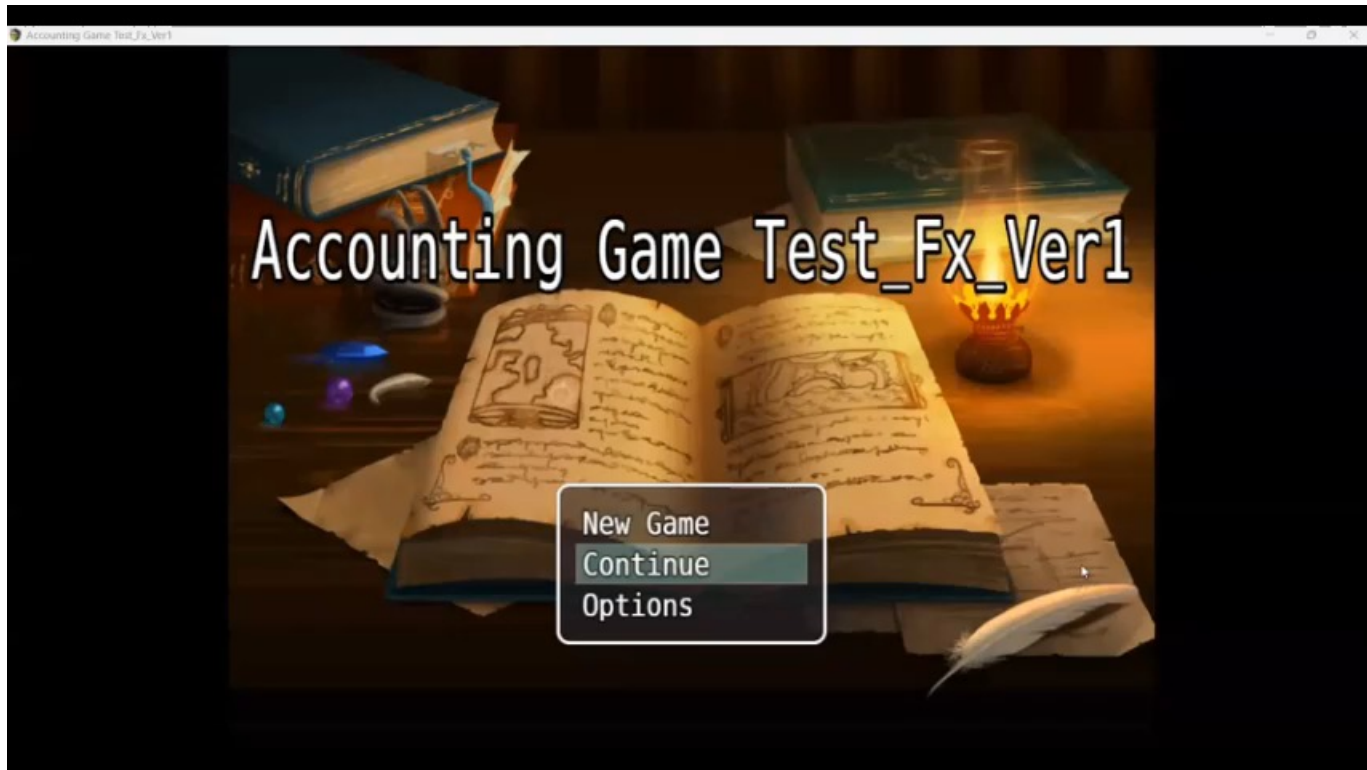
- In this study, the authors first built a 20- 30min RPG for "introduction to financial accounting."
- After the development of the game, the authors will introduce the game to accounting classes and test the impact of the RPG in accounting gamification on students' motivation and flow. Inspired by Sanchez et al, (2017), we also add "ludicization" elements in the game to increase the
- playfulness.

Game background

- The authors designed the game by using RPG Maker MV. The background of the story is a Rutgers student who comes to a mysterious town. In the town, the student meets Prof. Xiong, who says he/her gets a bakery as a lottery reward. The bakery needs help with bookkeeping.

Design

Game Demo



Design

Hypotheses Design

- Based on Zhang's et al. (2021) literature review on RPG education, we found RPG would most impact students' achievement and motivation. In this study, we'd like to focus on motivation.
- H1: RPG with an integrated case study has a positive impact on motivation.
- H2: RPG with an integrated case study has a positive impact on flow.
- H3: RPG with an integrated case study has a positive impact on attitude.

Questions Remains

- Should we also test performance?
- Is there a way to emphasize "integrated" story more in the experimental design?
- Although no such RPG research appears in accounting, many other subjects have proven the above hypotheses. How can the article be more distinguished?
- How about test the "ludicization" surprises in the game?

Survey and experiment Design

- The survey and experiment are under design.

Comments?



The background of the slide features a large, faint watermark of the Rutgers University seal. The seal is circular with a sunburst in the center and the words "RUTGERS UNIVERSITY" around the perimeter.

RUTGERS

THE STATE UNIVERSITY
OF NEW JERSEY

The Usefulness of Large Language Models in External Audits

Norah Xu

Kevin Moffitt

Objectives

- The primary objective of this research is to investigate the potential applications of LLMs in external audits and to evaluate their usefulness in enhancing the audit process.
- This research will evaluate LLMs in the context of the stages of the audit and the typical tasks conducted in each stage.

The stages of audit	How LLMs conducted	How to evaluate
Planning	<ul style="list-style-type: none"> Comprehend and summarize relevant documents such as financial reports, prior audit reports, and industry news. Draft the audit plan 	<ul style="list-style-type: none"> Compare the LLM's summaries and draft plans with those made by human auditors
Risk Assessment	<ul style="list-style-type: none"> Analyze text from various sources to identify potential risks. 	<ul style="list-style-type: none"> Compare the LLM's identified risks with those identified by human auditors
Internal Controls Testing	<ul style="list-style-type: none"> understand and explain control processes based on descriptions in the auditee's documentation. 	<ul style="list-style-type: none"> Compare the LLM's understanding with that of a human auditor
Substantive Testing	<ul style="list-style-type: none"> Not be directly involved 	
Reporting	<ul style="list-style-type: none"> Draft the report 	<ul style="list-style-type: none"> Evaluated based on the clarity, accuracy, and completeness of the LLM's draft compared to a human-generated report.
Follow-up	<ul style="list-style-type: none"> Review updated documents and comparing them with previous versions. 	<ul style="list-style-type: none"> The LLM's ability to identify changes.

Research progress

- Literature review
 - Research LLMs and their current applications;
 - Research current AI Applications in Auditing;
 - Research potential applications of LLMs in Auditing;
 - Research metrics for evaluating LLM applications in auditing.
- Methodology
 - Develop a framework for identifying and evaluating LLM applications in the external audit.
 - Apply this framework to the Planning stage and its associated tasks for a sample company.
 - Solicit feedback from audit professionals—is the framework useful and are the initial results informative?

Research progress (cont.)

- Results
 - Continue to fine tune the framework and apply it to the remaining stages of the audit.
 - Document the results and discuss findings.
- Limitations, Conclusion, Introduction, Abstract



The background of the slide features a large, faint watermark of the Rutgers University seal. The seal is circular with a sunburst in the center and the words "RUTGERS THE STATE UNIVERSITY OF NEW JERSEY" around the perimeter.

RUTGERS

THE STATE UNIVERSITY
OF NEW JERSEY

Bad Debt In Emergency Department

Yao Yao

Background

- In 2021, the amount spent on health care in the United States increased by 2.7 percent to \$4.3 trillion, or \$12,914 per person. Health spending accounted for 18.3 percent of the nation's GDP.
- One of the significant reasons for the rising cost is unpaid medical bills / medical bad debt (D. Shi, J. Zurada and J. Guan 2015). In addition, hospitals usually end up paying 30%-50% of recovered bad-debt revenue to collection agencies.
- Furthermore, uninsured, underinsured and patients with high-deductible plans add further pressure on the bad debt issue.
- According to Durand-Zaleski I et.al (1993), the surgical department has the strongest correlation to bad debt. However, the emergency department also plays a significant roles.

Emergency Department

- The emergency department is at the center of many significant issues in healthcare today. Previously the emergency department was referred to as a "loss leader" for revenue in many hospitals
- One of the major differences between emergency departments is managing the requirements for EMTALA — the Emergency Medical Treatment and Labor Act. Enacted by the U.S. Congress in 1986, it requires facilities to provide medical screening and stabilization to all patients, regardless of insurance status or ability to pay
- EMTALA also requires any hospital receiving Medicare payments to:
 - Provide incoming patients with a medical screening examination (MSE)
 - Stabilize patients who have an emergency medical condition
 - Transfer patients or accept them as appropriate to the care their facility can provide



Triage

Motivations

- Bad debt prediction is a complex task, and it is extremely difficult in the healthcare context. Unlike financial institution, health providers are unable to obtain any financial information due to legal and moral reasons, which forces hospitals to act as unintentional lenders.
- There is very little research exploring machine learning techniques in predicting bad debt in the healthcare industry.
- Despite the fact that many studies concentrate on the Emergency Department unit, they are more focused on predictions of either ED overcrowding or Triage outcomes leveraging machine learning.

Literatures	Variables	Models	Data Source
Zollinger et al.(1991)	Total hospital charge; Total hospital revenue; Patient information (insurance, employment, marital, gender, diagnosis, discharge status)	Multiple Regression model	985 patients from 28 Indiana hospitals
Veletsos (2003)	credit factors, demographic information, and previous organizational payment patterns	IBM Intelligent Miner and DB2	2400 patients (inpatients & outpatients) from Florida hospital
Zurada et.al (2005)	Patient Age (PA), Patient Gender (PG), Injury Diagnosis Code (IDC), Dollar Amount of the Claim (DAC).	Neural networks, decision trees, logistic regression, memory-based reasoning, the ensemble model	6180 cases from a healthcare company
D. Shi et al.(2016)	Days, Debt, Patient information (Gender, Age, Marriage, Occupation, insurance), Ngrade (level of nurse care), Injury diagnosis code (IDC), Total and other subtotal costs,	Bayesian networks, decision trees, KNN, random tree, random forest, radial basis function neural networks, bootstrap aggregating and neural networks	179 cases from China hospital
Liu N et al. (2022)	ED triage, ED disposition, inpatient discharge, and after inpatient discharge, patient information, ED waiting time, ED boarding time, patient loading time, time	Machine learning based AutoScore (Interpretable SERT); Random forest, Regression model, deep learning	EHR Dataset from Singapore General hospital over 13 years

Data Collection and Model Developments (Help Needed)

Variables

- Patient Information (Age, Gender, Insurance)
- EHR dataset (ED): Triage Information (Triage Class, Disposition, Diagnosis code, Outcomes)
- The total amount of claims
- **Good/Bad/Unknown**

