

# DISCOVERING NON-TRADITIONAL DATA SOURCES FOR BUSINESS INNOVATION

Harsimrat Pandher (VT), David Park (VT), Daniel Wilkin (VT), Joseph Kim (VT)

Claire Kelling (PSU) with Gizem Korkmaz and Stephanie Shipp (SDAL)

Sponsor: Gary Anderson, The National Center for Science & Engineering Statistics at NSF

## Project Description

- The National Center for Science & Engineering Statistics at NSF traditionally measures innovation through surveys of selected companies (e.g., Business R&D and Innovation Survey (BRDIS)). While BRDIS measures innovation incidence, i.e., the number of innovating firms, NCSES is interested in exploring the possibility of using non-traditional data to richer and complementary innovation measures.
- This project focuses on identifying non-traditional data sources to measure product innovation. Non-traditional data sources include administrative and opportunity data (e.g., product announcements, press releases, social media) that are obtained through databases, web-scraping or queries of selected company websites. The goal is to assess the feasibility of using these data sources and to develop methods to measure business innovation.

## Data Discovery

- The data discovery process aims to identify particular innovations, and where those innovations have left behind observable and traceable evidence that can then be captured with techniques such as web-scraping and text mining.
- We identify innovations using two approaches: (i) awards offered to particular innovations, and (ii) company deep dives. Additionally, we conduct a literature review to identify data sources prior researchers have employed.

### Company Deep Dives

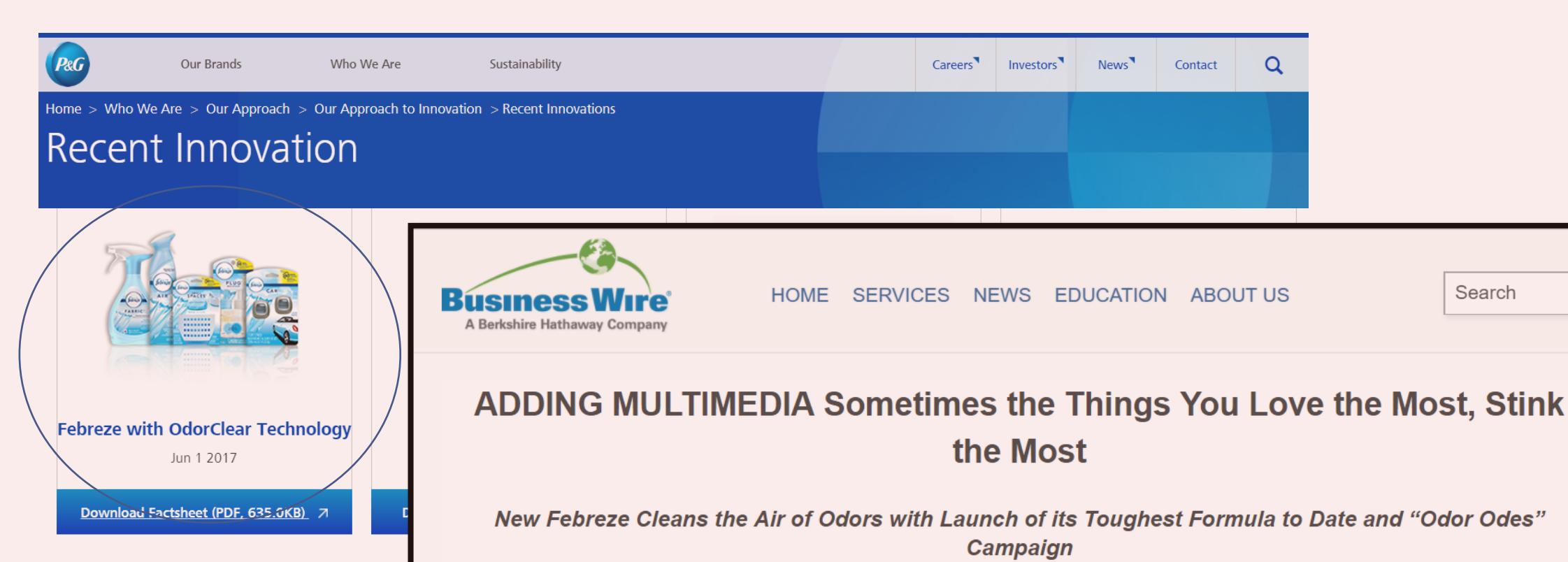
Company deep dives thoroughly examine *public sources of data* to identify particular innovations and identify data sources that might be used to identify initial market entry of those innovations.

#### Criteria for choosing companies:

- Size
- Industry
- Connections (for future interviews)

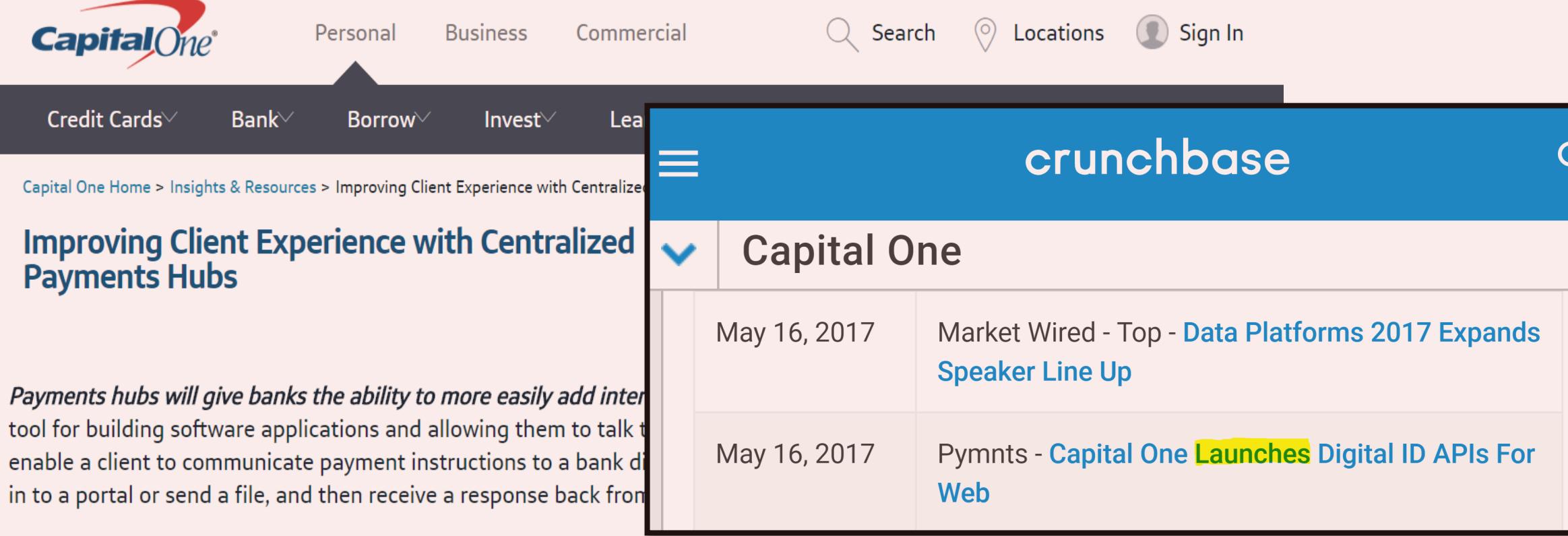
#### PROCTER & GAMBLE:

- “Innovation starts with the consumer”
- Open Innovation: Connect + Develop
- Product & Marketing Innovation (e.g., Tide Pods)



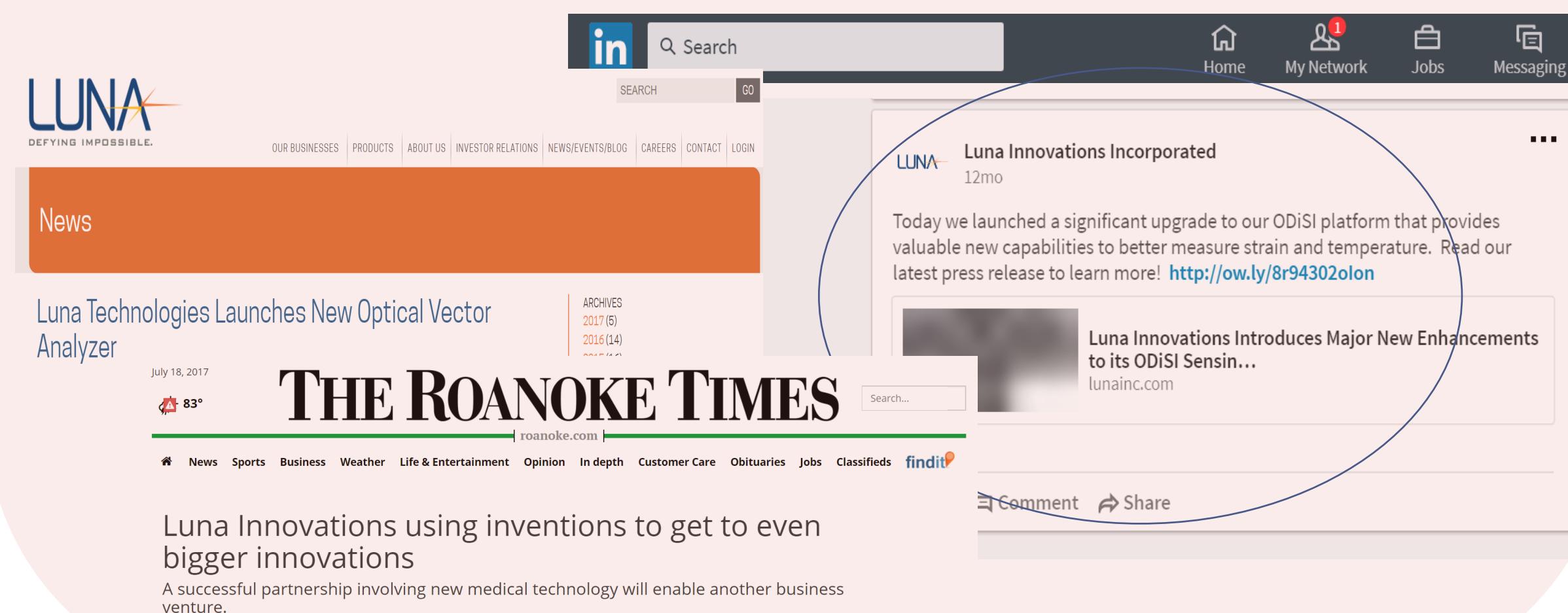
#### CAPITAL ONE:

- “To design new experiences that empower customers to better manage their finances”
- ENO - Capital One assistant text message
- Capital One Labs - towards the digital banking

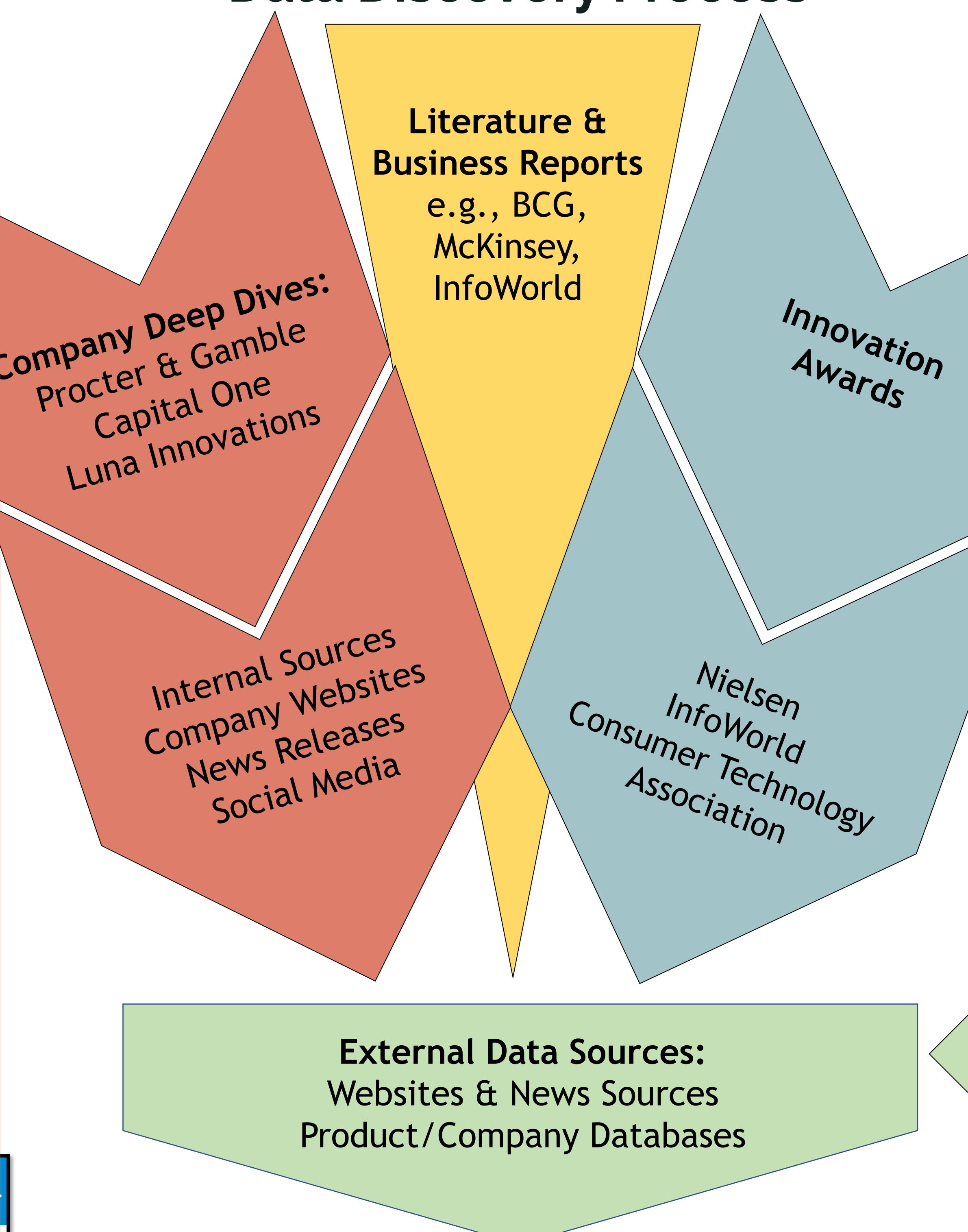


#### LUNA INNOVATIONS:

- “Research + Development”
- Product Development (e.g., Fiber Optic Sensor)
- Create subsidiaries for some successful products they offer

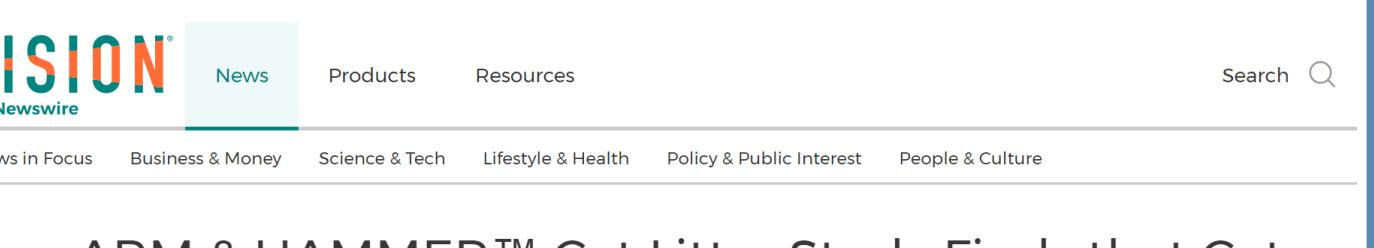
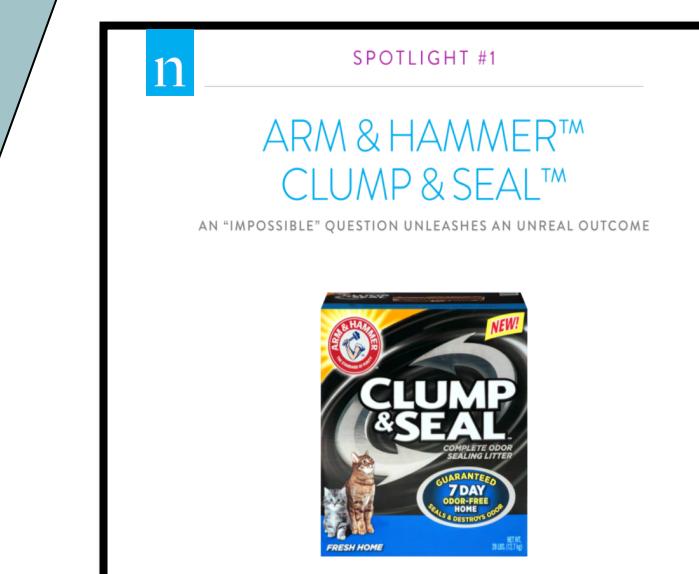


### Data Discovery Process



### Innovation Awards

Selected Awards	Product Awarded	Company	Industry
Nielsen	Arm & Hammer Clump & Seal	Church & Dwight Co.	Consumer Goods
	Bai	Bai	
InfoWorld	Amazon Alexa	Amazon	Technology
	TensorFlow	TensorFlow	
Consumer Technology Association	Aipoly Vision	Aipoly	Technology
	Cota Tile	Ossia	



#### Criteria used to rate websites/databases:

- Does it mention a launch?
- Does it have a specific launch date?
- Was it released at the time of the launch?
- Can the previous criteria be applied to other products (scalability)?
- How accessible is it?

SELECTED WEBSITES	Industry	DATABASES	Information Provided
Mass Device, Medical Design & Outsourcing	Medical	CrunchBase	Company Information, News, Social Media Links, Product Releases
Food Business News, Food Processing	Food	AngelList	Company Info, News, Social Media Links, Product Releases
Ars Technica, The Verge, Inc., Digital Trends, Engadget, Mashable, The Next Web	Science & Technology	Owler	Company Information & Analysis, News
		VentureBeat	Technology Innovation, News Reports
		F6S	Startups, Contests, Investment Info.
		Mattermark	Company Information & Analysis
Business Wire, Business Insider, Wall Street Journal, Media Post, PR Newswire, Kickstarter, Product Hunt, Reuters	Many Industries	Dow Jones VentureSource	Company Analysis, Investor Info.
		CB Insights	Company Analysis, Patents, Startups, Investor Info., News, Social Media
		Gust	Early-Stage Investments, Company Info.

Future work involves: • identifying more websites and databases • capturing other data sources such as trade magazines, SEC filings  
• conducting company interviews to identify data sources and validate findings with follow up interviews