

# **DATA JOURNALISM TO BRIDGE THE INFORMATION ASYMMETRY**

**Dr. Uma Shankar Pandey**



# JOURNALISM

- Journalism makes claims that are both incredibly modest (it simply reports the news of the day, quickly) and incredibly wide ranging (it possesses methodological techniques that allow it to objectively parse social reality)

*\*Anderson, C.W.. Apostles of Certainty (p. 3).*

# TRUTH CLAIMS: NARRATIVE JOURNALISM

- The Narrative Journalism form used tools of literary fiction, including exquisite detail, internal monologue, short story structure with a character, a problem and a resolution.

# PROBLEMS WITH NARRATIVE JOURNALISM

- It soon became obvious that narrative journalism tempted many writers to avoid fact-checking for fear that the facts might ruin a good story.

# PRECISION JOURNALISM

- The book by Philip Meyer published in 1973.
- Precision journalism treated journalism as if it were a science, adopting the scientific method, scientific objectivity and scientific ideals to the process of mass communication.

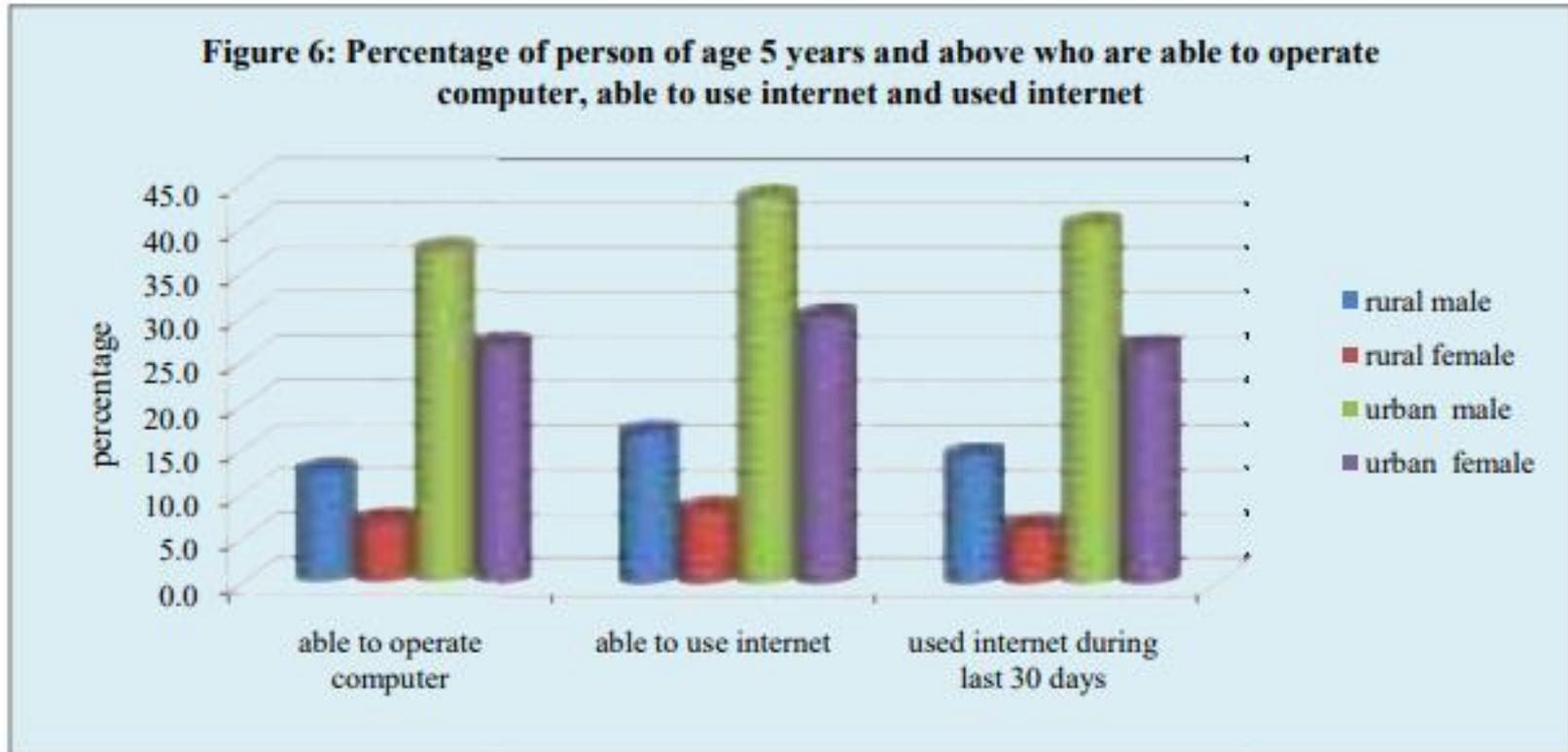
# PRECISION JOURNALISM QUESTIONS

- How to find information?
- How to evaluate and analyse it?
- How to communicate it in way that will make sense through the information overload and reach the people who need and want it?
- How to determine the amount of precision needed for a particular story?

# THE PRESENT DIGITAL ECOSYSTEM

- **There's 40 trillion gigabytes of data (40 zettabytes) in the world at present.**
- **If you were to store 40 zettabytes on DVDs, your stack of DVDs would be long enough to circle Earth 50 times.**

# PEOPLE WITH INTERNET FACILITY



- \*Key Indicators of Household Social Consumption on Education in India, NSO 2018 \*\*Smartphones not considered as computer

# INFORMATION ASYMMETRY

- The inability to take in and process it with the speed and volume that it comes to us—is one of the most significant problems that citizens face in making choices about how to live their lives.

# EVERYDAY DATA STORIES

- Data journalism allows us to go beyond the anecdote and provide evidence that something is happening, and we can show you what it is and why.

# ADDED VALUE FOR READERS

- Using data, the job of journalists shifts its main focus from being the fastest to report the stories to being the ones telling us what a certain development might actually mean. To decipher the signal from the noise becomes an important task

# HOW TO GET DATA

- In a regular Google Search you can start by looking only for spreadsheets (by appending your search with 'filetype:XLS or filetype:CSV')

# PUBLIC DATA

- **RBI Database Of Indian Economy**
- **Ministry Of Statistics And Programme Implementation Dataset**
- **Gateway To Indian Earth Observation**
- **National Portal of India**
- **Survey Of India**
- **India Weather Data**
- **Import Exports Datasets**
- **Wildlife Institute of India Dataset**

# DATA FROM THE WEB

- Get data from web-based APIs, such as interfaces provided by online databases and social media sites including Twitter, Facebook and Instagram.
- With Web scraping you extracting structured content with small pieces of code written in a programming language such as Python.
- Dataminer
- Import.io
- Outwit
- Downthemall.

# IMPORTANT DATA SOURCES

- Google Public Data Explorer
- Google Dataset Search
- Global Forest Watch
- DataPortals.com
- Election Databot (US Elections)
- Google Trends
- The International Consortium of Investigative Journalists (ICIJ)
- ProPublica's Data Store

# IMPORTING DATA INTO SPREADSHEETS

- Google Spreadsheets
- =IMPORTDATA("http://www.census.gov/2010census/csv/pop\_change.csv")
- =IMPORTHTML("https://www.mohfw.gov.in", "table", 1)
- =IMPORTFEED("https://www.thehindu.com/news/national/?service=rss", "", true, 100)

# LET'S DOWNLOAD SOME DATA

# DATASET STORIES BY PAUL BRADSHAW

- **Top of the pops/flops**
- **‘Holding up a mirror’**
- **A freak occurrence/unusual change**
- **A claim debunked or supported**
- **Trend**
- **Postcode lottery**
- **“Juking the stats”/“No accountability”**
- **Victims, Winners and Losers**
- **Things happen together**
- **“It’s not working**

*\*Data Journalism Heist by Paul Bradshaw*

# AN EXAMPLE

NATIONAL

## Coronavirus | 18 States/UTs have recovery rate higher than national average of 64.2%: Health Ministry



SPECIAL CORRESPONDENT

NEW DELHI , JULY 10, 2020 16:16 IST

UPDATED: JULY 10, 2020 16:30 IST

SHAR

COVID-19

807283

REPORTED CASES

[VIEW MAP](#)



PRINT

A

A

A



# LET'S ANALYZE COVID19 DATA

- <https://ourworldindata.org/coronavirus-data>

# ALTERNATIVES

- Think of this as an exercise in imagination. Be creative by thinking of the alternative stories that might be consistent with the data and explain it better, then test them against more evidence.
- ‘What other story could explain this?’

# PITFALLS

- There's so much data about that you will often be able to find confirmation of your prior beliefs if you simply look around a bit.
- Data journalism, adds little value if you are not open-minded. It is only as objective as you strive to make it, and not by virtue of being based on numbers.



# FLOURISH STUDIO

