SEVEN CORE SKILLS
of DATA STORYTELLING

- IDENTIFY
- OBTAIN
- EXTRACT
- CLEAN
- ANALYZE
- VISUALIZE
- TELL A STORY

SEVEN CORE SKILLS *of* DATA STORYTELLING

Data-driven stories can seem magical, giving professionals a superpower of persuasion. But showing workers how to create such stories can be a painful process. Follow these tips and tricks taught at Boston University's Data + Narrative workshops, and you'll discover that teaching data-driven techniques need not drive you crazy.

IDENTIFY

OBTAIN

EXTRACT

CLEAN

ANALYZE

VISUALIZE

TELL A STORY

1. Go beyond Google and gather more data than you think you need. Extra data can give people a richer understanding of an issue. Numbers can be used for bench-marking or add an unexpected dimension to a story.

2. Get the data you want, often for free. Encourage students to explore data that might be most relevant to their work, including the 300,000 databases filled with government info at data.gov and statistics about 167 million companies at opencorporates.com.

3. Use the latest tools to extract data from PDFS and other documents. Tools like Cometdocs and Tabula will simplify tedious and aggravating work.

4. It's essential to make sure your datasets aren't filled with duplicate or inaccurate information. Use tools that can standardize data in seconds and consider keeping a data diary, recording every step you take to locate and manipulate data.

5. When it's time to examine data, think first about key points you may want to make. Search for patterns. Combine and compare datasets. It's easy to make lots of pivot tables to find outliers.

6. Choose what type of graphic will best display your data. Beware of pie charts, which often create confusion. Better to rely on simpler bar charts and graphs.

7. Use simple words, strong verbs, and consider a narrative structure, complete with a beginning, middle and end. Take your audience on a journey to solve a business problem, and think about ways to engage your audience's emotions.

To learn more about and register for Boston University's June workshop, click here.

Contents

This ebook is brought to you by Boston University's College of Communication and its Data+Narrative Workshops. To register for our June workshops, please **click here**. For more information, contact Mark Krumm, training manager, at **mkrumm@bu.edu**.

Articles by Rochelle Sharpe
storytellingsolutions.com

Illustration and Design by Jaclyn Cimino
Boston University College of Communication
bu.edu/com

Data+Narrative
Data-Driven Storytelling Workshops
combeyond.bu.edu/workshop/data-narrative

Copyright 2019

The Seven Core Skills of Data Storytelling

By [Rochelle Sharpe](#)

Data-driven stories can seem magical, giving professionals a super-power of persuasion.

But training people how to create such articles or presentations can be a daunting task. There are lots of hurdles to overcome: dealing with data that's out of date, getting lost in mounds of statistics, or being forced into such mind-numbing chores as manually pulling data into spreadsheets. Discovering duplicate numbers, holes in data, or even disinformation can make almost everyone nostalgic for the good old days of working with just pencil and paper.

Fortunately, there are lots of tricks and tools to make the training process less agonizing. **Some advice offered at Boston University's storytelling workshops** may help show that teaching data-driven story techniques need not drive you crazy.

Gather More Data

Professionals may think they need only one set of data to tell their stories. But often, extra data can give people a richer understanding of an issue. There may be numbers that can be used for benchmarking purposes or add an unexpected dimension to a story. Since we really don't know what we don't know, it's worth urging your trainees to spend the time searchingfor extra numbers – especially in this era where new data sets pop up almost every day.

Where to hunt for more data? Urge students to explore specific websites that might be most relevant to their work. Are they interested in government information? They may want to check out **data.gov**, where they can easily access more than 300,000 databases. Do they want information about corporations? They should look at **opencorporates.com**, which provides information about more than 167 million companies around the world.

Extra data can give people a richer understanding of an issue.

Make sure your trainees know how to use advanced searches on Google, limiting queries to specific sites with the "site:" filter. They can also experiment with the beta "Google Data Set" search, where users can plug in a topic to search for data.

Time-Saving Tools

Once trainees find the data they want, they'll need to collect and organize it. To simplify what could be a tedious or aggravating process, consider teaching tools like Tabula, which extracts data from PDFs, or Cometdocs, which converts PDFs to different file formats for easy

analysis. Do they need to pull data from multiple spreadsheets together? There's a tool for that, too: VLookup, a versatile function in Excel that allows you to retrieve data from a table.

To make sure the dataset is accurate and not filled with lots of duplicate entries, it's essential that trainees learn to clean their data. There are several tools that can help standardize data in seconds. Teaching
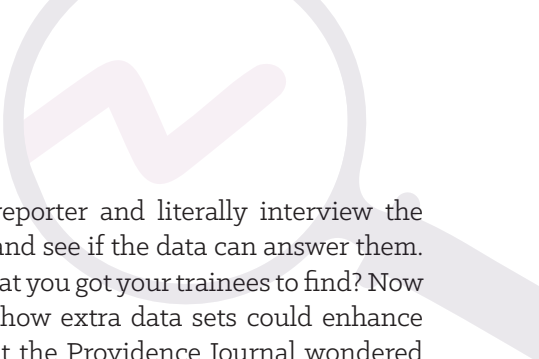
Creating a data diary can be invaluable.

the need to create a data diary would also be wise. It may take time for professionals to record every single step they take to locate and manipulate their data. But if they ever need to figure out whether they've somehow inserted errors in their work, the diary can be invaluable.

Elevate the Analysis

When analyzing data, trainees must keep in mind their end goal: a data-driven story. Whether writing an article or a preparing a PowerPoint presentation, they should look at the data first to figure out the key points they may want to make. Search for patterns. Combine and compare data sets. It's easy to make lots of pivot tables in Excel and spot outliers. Only after trends are identified should professionals start doing deep dives into the data and look for intriguing nuances.
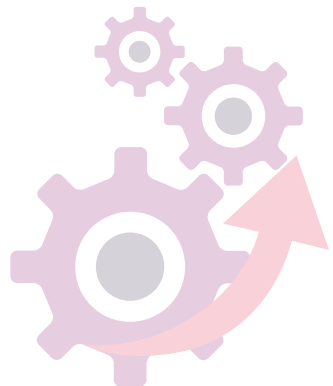
Often, it's helpful to act like a reporter and literally interview the data. Set up a series of questions and see if the data can answer them. Remember those extra data sets that you got your trainees to find? Now is the time they should consider how extra data sets could enhance their analysis. When journalists at the Providence Journal wondered years ago why a string of kids were being injured in bus accidents, they didn't just get a list of school bus accidents. They also got lists of all school bus drivers, all drivers' license violations, and all felons in the state. When they were done comparing and contrasting these lists, they discovered school bus drivers who had been involved in more than 20 accidents during a three-year period as well as many who were drug dealers and convicted felons.

Teach students when they should act like misers.

Story Time: Less Really Is More

When it's time to finally **prepare the story**, urge your students to change their behavior. Get them to reverse course completely. Until now, they've learned to be in a kind of spendthrift mode, going on sweeping sprees to find and collect reams of data and then analyzing what they've found in as many ways possible. Now, it's time for them to act like misers, hoarding most of the information and showing only what's absolutely necessary. This is not an act of selfishness. The audience really doesn't want to know all the gory details. They don't want to slog through reams of charts and graphs.

Your trainees may not want to believe this. Their first impulse will be to show everyone all the work they've done and all the cool information they've discovered. But this is the time for self-restraint. After making sure their data is accurate, their top goal: don't lose the audience. They've done all this great work. Don't let people ignore it.

Get your students to convey their ideas simply and in engaging ways. To do that, they'll need to do the hard work of figuring out what matters most. Don't let them skip this process, even if it takes some time. Remember Mark Twain's famous quip: "I didn't have time to write a short letter, so I wrote a long one instead."

Your happy ending: a compelling story that persuades the powers that be.

Get them to consider telling their story in a narrative form, complete with a beginning, middle and end. They might want to start with the business problem, then take their audiences on a journey about people's efforts to tackle the problem, and finally end with how the problem can be resolved. As they tell their tales, they should think of ways to engage their audience's emotions.And make sure they cut out any clutter – in both writing and data visualizations. Clutter can only distract, confuse – and ultimately -- bore the audience.

With data visualizations, make sure they simplify their data and put it in an understandable context. Urge them to avoid pie charts. Those often create confusion. Instead, they should rely on simpler bar charts and graphs. If they need to show geographical detail, they could try a tool like Carto, which lets them map data. To show changes over many years, they could consider timelines. Once they master the basics, they can think about adding some wow factors, like interactive visualizations. They can also make charts with multiple layers, which allow people to drill down into a graphic to get more details about a project.

Putting together such stories may seem daunting at first. But by following some simple steps, your trainees can reach your happy ending: a compelling story that persuades the powers that be.

Boston University's Data+Narrative: Data-Driven Storytelling Workshops are held the first two weeks of June at BU's campus in Kenmore Square in Boston and online for remote learners.

The Power of Story

*By **Rochelle Sharpe***

As a journalist, I see the world through stories.

I'm always on the lookout for great ideas and often collect articles that are beautifully written. Sometimes, I analyze stories that I love and try to figure out what makes them so wonderful. Why are they so memorable? What gives them such power?

My journalism career offers some clues. Over the years, I've covered presidential campaigns, done a lot of investigative reporting, written all kinds of quirky features, and yes, even won a Pulitzer Prize. But what's really surprised me is my one story that continues to have staying power.

What gives stories their power?

I wrote this story for The Wall Street Journal more than 20 years ago, several years before newspapers had gone online and long before the term "going viral" even existed. But people still track me down, sometimes whispering into the phone when they call to talk to me about it.

The story was about a woman with severe body odor, the kind that literally made people gag. The pungent odor came intermittently and without notice, and the woman could not control it. She couldn't even detect it because a bad car accident had robbed her of her sense of smell. Still, the odor became a such distraction in her office that it eventually cost her her job. She took her company to court, arguing that it was violating the Americans with Disabilities Act. But she lost her case.

Certainly, this story had its sensational elements, which fueled a bit of a media frenzy at the time. The woman appeared on Oprah, got mentioned on David Letterman's late night comedy show, and even wound up on the cover of the National Enquirer. But the phone calls over the past 20 years had nothing to do sensationalism. They were from people who also suffered from body odor, who wanted to speak with this woman and commiserate with her. To them, this woman had broken an important taboo. She was willing to expose her vulnerabilities and they wanted her to know that she was not alone. Their actions reminded me of the provocative title of an old collection of Esquire Magazine essays: "All Our Secrets are the Same."

These days, when I teach writing workshops, I always talk about the power of emotion in stories. I like to invoke the work of James K.A. Smith, who suggested in his book "You Are What You Love" that to really connect with others, we can't just behave like people are "brains on a stick."

Stories are just data with a soul.

Brene Brown, the dynamic professor who studies shame, courage, and vulnerability, makes a similar point. The author of multiple best sellers, Brown is probably best known for delivering one of the world's most popular TED talks, "The Power of Vulnerability," which has been viewed more than 35 million times. She argues that the desire for connection often gets derailed by shame and fear. If people embrace -- rather than hide -- their vulnerabilities, they can live more authentic lives and experience more joy, creativity, and love.

At the start of the TED talk, Brown describes her discomfort with an event planner who wants to call her a storyteller, a name that she worries makes her sound like a flaky academic. But in the end, Brown welcomes the moniker and provides an intriguing definition of story, too. "I'm a qualitative researcher," she says. "I collect stories ... Maybe stories are just data with a soul."

Yes. Data with a soul. That is a useful construct for the legions of companies trying to incorporate storytelling into their businesses. In this era of big data, all kinds of firms are searching for ways to present mounds of information, and many are discovering that stories can help them make sense of their worlds.

We can't just behave like people are "brains on a stick."

That's why this ancient form of expression is experiencing such a revival these days. In our high-tech, fast-paced society, where disconnection is on the rise, stories can have more power than ever.

---

*Rochelle Sharpe will be offering a **satellite workshop** on writing at Boston University's Data + Narrative: Data-Driven Storytelling Workshops during first two weeks of June. Please visit our workshop site to learn more.*

THE POWER OF DATA-DRIVEN STORYTELLING: BOSTON UNIVERSITY'S DATA+NARRATIVE WORKSHOP

Boston University's College of Communication offers the definitive workshop to perfect your data-driven storytelling skills.

Learn how to tell data stories with Pulitzer Prize-winning journalists, a CNN data analyst, and other experts at Boston University's fifth annual Data + Narrative workshop. Studying data analysis, narrative writing, and data visualization will position you to engage your audience through powerful data-driven stories.

In today's data-flooded environment, you need the skills to decipher information and craft compelling stories for multiple audiences. Let the faculty and guest instructors at the Data + Narrative Workshop teach you how to make compelling presentations to influence your audience. Our three tracks of instruction and small class sizes insure that you will learn at the level best suited to your experience. You'll also have the opportunity to take an extra writing workshop with a Pulitzer Prize-winning journalist.

Who Should Attend?

Workshop attendees gain a practical tool kit of data and digital storytelling skills to build on year after year. Students have included persons in philanthropy, law, education, social work, public health, library sciences, engineering, journalism, finance, real estate, retail, marketing and public relations, federal, state and municipal government, and more.

Whether to educate and inform your audience, win a key argument, tap into key funding sources or edge out the competition for new business, our workshops are your guide to success.

On Campus: June 3 – 7, 2019
Online: June 10 – 14, 2019

This year, you may choose to attend the weeklong workshops either on the Boston University campus or join us online in a robust virtual learning environment.

Learn More

Visit our **website** to register for our course and supplemental writing workshop, and to reserve your seat in the sessions of your choice. Meet faculty, learn more, and ask questions in our online **Infosession** April 16th, Noon EDT.

WHY DATA+NARRATIVE?

Optimize the power of your data to more effectively persuade your audience, inform and educate employees, clients, customers and shareholders, disrupt the competition, harness new funding sources or win a pivotal argument.

Data fluency is a must-have skill set in today's digital workplace. Success starts with knowing how to find the meaning in every dataset – regardless of its size or format.

Our intensive 5-day workshops equip you with a practical, ready-to-go toolbox of knowledge, techniques and strategies to become a master data storyteller – no matter your experience or skill level.

Deft storytelling skills are needed to understand and interpret today's endless amount of digitized professional and personal data. Those skills are far more sophisticated than looking at a prepared dataset and creating an infographic.

What can you do with a data story?

- Collect, analyze and present data to spur your audience to act

- Showcase key data interactively to outline your goals

- Incorporate persuasive data findings in presentations to convince stakeholders

- Stimulate audience engagement by including compelling data in external communications and social media

- Build stakeholder buy-in and brand advocacy with data-rich communications

- Identify the most persuasive statistics to secure a legal victory

- Extract and standardize data and create outstanding data visualizations to convey a trend or fact-based conclusion

WORKSHOP INFO

Dates

Session I (*on campus*):
June 3 - 7, 2019

Session II (*online only*):
June 10-14, 2019

Fees

For Professionals:	$2,150
For Academics, Non-Profits, and Government Entities:	$1,600
For BU Alumni & Affiliates:	$1,250
For Students:	$650
Satellite Session (Writing with Data):	$500

SPECIAL SATELLITE WORKSHOP:
Writing *with* Data

In this era of high-tech distractions, when humans have shorter attention spans than goldfish, how can you cut through the cacophony and get your message heard? The age-old craft of storytelling can be a potent tool.

Come to this satellite workshop to discover the power of narrative writing and learn how to make your prose clearer, more evocative, and surprisingly persuasive. This half-day course will be taught by Rochelle Sharpe, a Pulitzer Prize-winning journalist, writing coach, and former staff writer for The Wall Street Journal. She will customize this workshop to meet participants' needs.

Participants will learn how to develop a storytelling mindset and the nuts and bolts of narrative techniques. They'll discuss word choice, sentence composition, easy ways to organize material, and how to create narrative structures. They also will consider best practices for incorporating data into prose and ways to make even the most complicated material simple and understandable. Along the way, Rochelle will demystify the writing process, showing how to make it less painful -- and maybe even fun.

Students will look all kinds of writing examples, ranging from passages from the Iliad to speeches by Donald Trump. They also participate in multiple short drills to practice what they are learning.

Feel free to bring in an article that you may want to discuss.

WORKSHOP **SCHEDULE**

DATA+NARRATIVE
Data-Driven Storytelling Workshops

This schedule is our Level 3 Training, **Data + Narrative with R**

SUNDAY, JUNE 2
Optional Session: Writing with Data

2:00 – 6:00PM
Satellite Session: Writing with Data
Instructor: Rochelle Sharpe, Pulitzer Prize Winning Freelance Journalist.
Learning Objective: Using Data To Form A Storyline
COM Room 209

MONDAY, JUNE 3
Day 1: Finding Data

8:00 – 9:00AM
Optional Software Installation/Technical Help Session
Come early if you need help reviewing software installation or other technical issues
COM Room 209

9:00 – 9:30AM
Data+Narrative: Getting Started
Instructor: Maggie Mulvihill, BU Data+Narrative Workshops Founder &
Executive Director, Associate Professor of the Practice in Computational Journalism
Learning Objective: Capture the big picture of data storytelling
COM Room 209

9:30 – 10:30AM
Extracting Reliable Data from Social Media: The Latest Tools
Learning Objective: Harness social media data to inform your story
COM Room 209

10:45 – 11:30AM
So Much Data: Finding High-Value Records for Your Project
Instructor: Margot Williams, Research Editor for Investigations, the Intercept
Learning Objective: Identify the best data for your project
COM Room 209

11:30 – 1:00PM
Lunch & Narrative Data Storytelling Writing Practice Exercises
Instructor: Rochelle Sharpe, Pulitzer Prize winning freelance reporter
Learning Objective: Transform numbers into narrative
COM Room 209

1:00 – 4:00PM
Afternoon Exercises, Story Practice Time & Project Selection
Learning Objective: Apply your data storytelling skills and decide on a project for the week
COM Room TBD

1:00 – 2:30PM
Optional Session: Deep Dive Into Key Data Resources
Instructor: Margot Williams, Research Editor for Investigations, the Intercept
Learning Objective: Isolate hidden data resources for your project
COM Room 209

2:45 – 4:00PM
Optional Session: Stats 101: Finding the Statistical Significance in Your Data
Instructor: Wayne Snyder, Boston University, Associate Professor of Computer Science
Learning Objective: Understand the statistical meaning in your data
COM Room 209

4:00 – 5:00PM
Project Challenges, Summary of Today's Lessons, Practice Exercises and Evaluations

5:00 – 8:00PM
Data Storytelling Happy Hour at Scoozi
All attendees, student coaches, staff and faculty welcome.
Scoozi is at 580 Commonwealth Avenue, Boston, MA 02215

TUESDAY, JUNE 4
Day 2: Extracting Data

9:00 – 10:30AM
Getting Your Data into R
Learning Objective: Use R data structures
Hariri Seminar Room, 111 Cummington Mall

10:45 – 12:30PM
Data Cleaning and Organizing
Learning Objective: Define tidy data and identify data that is not tidy
Hariri Seminar Room, 111 Cummington Mall

1:30 – 3:00PM
Exploratory Data Analysis
Instructor: Haviland Wright, Boston University, Professor of the Practice, Mathematics and Statistics
Learning Objective: Describe how EDA fits into data analytic strategies
Hariri Seminar Room, 111 Cummington Mall

3:15 – 4:45PM
Reproducibility
Learning Objective: Students produce markdown documents that describe their data prior to analysis
Hariri Seminar Room, 111 Cummington Mall

4:45 – 5:00PM
Project Challenges, Summary of Today's Lessons, Practice Exercises and Evaluations
Hariri Seminar Room, 111 Cummington Mall

7:00 – 9:00PM
Satellite Session: Writing with Data
Instructor: Rochelle Sharpe, Pulitzer Price Winning Freelance Journalist
Learning Objective: Using data to form a storyline
COM Room 209

WEDNESDAY, JUNE 5

9:00 – 10:30AM
Unsupervised Learning: Finding the Key Variables in Your Data
Learning Objective: Use principal component analysis to identify the variables in your data that matter most and could be potential stories
Hariri Seminar Room, 111 Cummington Mall

10:45 – 12:30PM
Supervised Learning: Using R to Build Decision Trees
Learning Objective: Use supervised learning to create decision trees in order to understand the relationships among variables in your data
Hariri Seminar Room, 111 Cummington Mall

1:30 – 3:00PM
Clustering to Identify Potential Stories in Your Data
Learning Objective: Use Kmeans and hierarchical clustering to group your data and better understand its significance and potential
Hariri Seminar Room, 111 Cummington Mall

3:15 – 4:45PM
The Power of Statistical Analysis
Instructor: Haviland Wright, Boston University, Professor of the Practice, Mathematics and Statistics
Learning Objective: Understand how to find the statistical significance of your data
Hariri Seminar Room, 111 Cummington Mall

4:45 – 5:00PM
Project Challenges, Summary of Today's Lessons, Practice Exercises and Evaluations
Hariri Seminar Room, 111 Cummington Mall

7:00 – 9:00PM
Satellite Session: Writing with Data
Instructor: Rochelle Sharpe, Pulitzer Price Winning Freelance Journalist
Learning Objective: Using data to form a storyline
COM Room 209

THURSDAY, JUNE 6
Day 4: Visualizing Your Data

9:00 – 10:30AM
Producing Static Presentations
Learning Objective: Use ggplot2, ggmap, markdown, and knitr to produce narratives in a static document
Hariri Seminar Room, 111 Cummington Mall

10:45 – 12:30PM
Producing Dynamic Presentations
Learning Objective: Produce dynamic slide presentations using markdown, ggvis, leaflet, knitr, and plotly
Hariri Seminar Room, 111 Cummington Mall

1:30 – 3:00PM
Making Web Pages with R
Learning Objective: Use Shiny to produce a webpage for your narrative
Hariri Seminar Room, 111 Cummington Mall

3:15 – 4:45PM
Summary and Review
Instructor: Haviland Wright, Boston University, Professor of the Practice, Mathematics and Statistics
Learning Objective: Integrate what you have learned in the previous sessions to produce a document, a presentation and a webpage
Hariri Seminar Room, 111 Cummington Mall

4:45 – 5:00PM
Summary of Today's Lessons, Practice Exercises and Evaluations
Hariri Seminar Room, 111 Cummington Mall

FRIDAY, JUNE 7
Day 5: Data Narrative: How to Tell Your Story

9:00 – 10:30AM
Data Storytelling: How To Spot Story Insights
Learning Objective: Tease out what you will need to convey before you hit send
COM Room 209

10:00 – 12:00PM
Data Narrative Studio
Learning Objective: Interact one-on-one with instructors and coaches on storytelling challenges
COM Room 209

12:00 – 1:00PM
Catered Lunch and Presentation: Framing the Narrative: How Data Forms Your Storyline
Learning Objective: Understand how to craft a powerful data-driven narrative
COM Room 209

1:00 – 3:00PM
Data+Narrative Lightning Showcase
Our finale Data+Narrative Storytelling Contest will feature attendees presenting their data stories; prizes to the top three data storytellers of the week!
Learning Objective: See how much we have achieved during this workshop
Hariri Seminar Room, 111 Cummington Mall

3:00 – 3:30PM
Final Wrap-Up and Evaluations
COM Room 209

1:15 – 4:00PM
Project Work, One-on-One Data Storytelling Coaching, Afternoon Exercises, Story Practice Time
Learning Objective: Apply your data storytelling skills and get help with your project

3:30 – 6:30PM
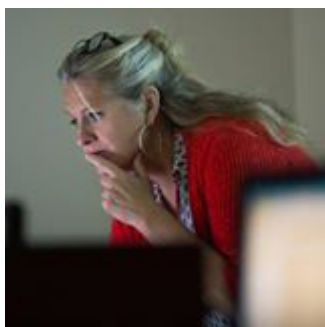### Optional Session: Writing with Data
Instructor: Rochelle Sharpe, Pulitzer Prize Winning Freelance Journalist
Learning Objective: Using data to form a storyline
COM Room 209

Instructor Biographies

Maggie Mulvihill

Maggie Mulvihill, a professor at the Boston University College of Communication, is the founder and director of Data + Narrative, the nation's first comprehensive university-based data storytelling workshop. An attorney, Maggie teaches data journalism at Boston University, leading students to produce major investigative projects. She co-founded the New England Center for Investigative Reporting and is also a Faculty Fellow at Boston University's Rafik B. Hariri Institute for Computing and Computational Science and Engineering. She serves on the Steering Committee of the Reporter's Committee for Freedom of the Press, the board of the New England First Amendment Coalition, and in 2014, was named to the Federal Freedom of Information Act Advisory Committee. She was a Nieman fellow in 2004-2005 at Harvard University.

Rochelle Sharpe

Rochelle Sharpe is a Pulitzer Prize-winning journalist with more than 25 years of experience. Now a freelance writer in Brookline, MA focusing on health and education issues, Rochelle has published stories in such publications as The New York Times and The Washington Post. Earlier in her career, she worked as a staff reporter for The Wall Street Journal, Business Week and USA Today, writing groundbreaking articles discussed at Congressional hearings, as well as on Oprah and the David Letterman show. A pioneer in computer-assisted reporting, Rochelle also has worked as a journalism professor, editor, and writing coach, and helped create an investigative reporting workshop for high school students at Boston University.