

COMPUTATIONAL JOURNALISM

Tuesday and Thursday Evenings, Spring of 2023
Instructors, Mark Hansen and Wayne Lee

JOURNALIST'S BRIEFING



You!

What experience do you have working with data, in or out of journalism?



Chang, Duojiao
dc3667



Cruz-Luviano, Alma
alc2313



Fan, Hangyu
hf2465



Galloway, Alana
amg2425



Jiang, Lai
lj2562



Marx, Sarah
sjm2255



Patwardhan, Anvita
ap4302



Shailer, Daniel
dls2226



Woods, Ellen
ehw2134



Wu, Ruoheng
rw2958



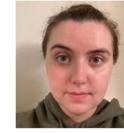
Brahmabhatt, Manan
mb4984



Fan, Yongyang
yf2635



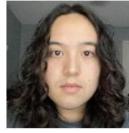
Gu, Jingyu
jg4536



Hanger, Rachel
rah2240



Hawkins, Diahmin
dlh2166



Lam, Daniel
dr12163



Li, Chongzhi
cl4324



Nguyen, Kieu-Giang
kn2521



Ngwenya, Mengezi
mn3040



Wei, Jiahao
jw4312



Zhou, Hongyan
hz2827

REVISED AND UPDATED 4TH EDITION

BILL KOVACH
& TOM ROSENSTIEL

— THE —
ELEMENTS
OF
JOURNALISM

What Newspeople Should Know
and the Public Should Expect

In their book [The Elements of Journalism](#), Bill Kovach and Tom Rosenstiel identify the essential principles and practices of journalism.

Here are 10 elements common to good journalism, drawn from the book.

Journalism's first obligation is to the truth

Good decision-making depends on people having reliable, accurate facts put in a meaningful context. Journalism does not pursue truth in an absolute or philosophical sense, but in a capacity that is more down to earth.

"All truths – even the laws of science – are subject to revision, but we operate by them in the meantime because they are necessary and they work," Kovach and Rosenstiel write in the book. Journalism, they continue, thus seeks "a practical and functional form of truth." **It is not the truth in the absolute or philosophical or scientific sense but rather a pursuit of "the truths by which we can operate on a day-to-day basis."**

This "journalistic truth" is a process that begins with the professional discipline of assembling and verifying facts. Then journalists try to convey a fair and reliable account of their meaning, subject to further investigation.

Journalists should be as transparent as possible about sources and methods so audiences can make their own assessment of the information. Even in a world of expanding voices, "getting it right" is the foundation upon which everything else is built – context, interpretation, comment, criticism, analysis and debate. The larger truth, over time, emerges from this forum.

As citizens encounter an ever-greater flow of data, they have more need – not less – for suppliers of information dedicated to finding and verifying the news and putting it in context.

Its first loyalty is to citizens

The publisher of journalism – whether a media corporation answering to advertisers and shareholders or a blogger with his own personal beliefs and priorities – must show an ultimate allegiance to citizens. **They must strive to put the public interest – and the truth – above their own self-interest or assumptions.**

A commitment to citizens is an implied covenant with the audience and a foundation of the journalistic business model – journalism provided “without fear or favor” is perceived to be more valuable than content from other information sources.

Commitment to citizens also means journalism should seek to present a representative picture of constituent groups in society. Ignoring certain citizens has the effect of disenfranchising them.

The theory underlying the modern news industry has been the belief that credibility builds a broad and loyal audience and that economic success follows in turn. **In that regard, the business people in a news organization also must nurture – not exploit – their allegiance to the audience ahead of other considerations.**

Technology may change but trust – when earned and nurtured – will endure.

Its essence is a discipline of verification

Journalists rely on a professional discipline for verifying information.

While there is no standardized code as such, every journalist uses certain methods to assess and test information to “get it right.”

Being impartial or neutral *is not* a core principle of journalism. **Because the journalist must make decisions, he or she is not and cannot be objective. But journalistic methods are objective.**

When the concept of objectivity originally evolved, it did not imply that journalists were free of bias. It called, rather, for **a consistent method of testing information – a transparent approach to evidence – precisely so that personal and cultural biases would not undermine the accuracy of the work.** The method is objective, not the journalist.

Seeking out multiple witnesses, disclosing as much as possible about sources, or asking various sides for comment, all signal such standards. This discipline of verification is what separates journalism from other forms of communication such as propaganda, advertising, fiction, or entertainment.

Its practitioners must maintain an independence from those they cover

Independence is a cornerstone of reliability.

On one level, it means not becoming seduced by sources, intimidated by power, or compromised by self-interest. **On a deeper level it speaks to an independence of spirit and an open-mindedness and intellectual curiosity that helps the journalist see beyond his or her own class or economic status, race, ethnicity, religion, gender or ego.**

Journalistic independence, write Kovach and Rosenstiel, is not neutrality. While editorialists and commentators are not neutral, the source of their credibility is still their accuracy, intellectual fairness and ability to inform – not their devotion to a certain group or outcome. In our independence, however, journalists must avoid straying into arrogance, elitism, isolation or nihilism.

It must serve as an independent monitor of power

Journalism has an unusual capacity to serve as watchdog over those whose power and position most affect citizens. **It may also offer voice to the voiceless.** Being an independent monitor of power means “watching over the powerful few in society on behalf of the many to guard against tyranny,” Kovach and Rosenstiel write.

The watchdog role is often misunderstood, even by journalists, to mean “afflict the comfortable.” While upsetting the applecart may certainly be a result of watchdog journalism, the concept as introduced in the mid-1600s was far less combative. Rather, it sought to redefine the role of the journalist from a passive stenographer to more **a curious observer who would “search out and discover the news.”**

The watchdog role also means more than simply monitoring government. “The earliest journalists,” write Kovach and Rosenstiel, “firmly established as a core principle their responsibility to examine unseen corners of society. The world they chronicled captured the imagination of a largely uninformed society, creating an immediate and enthusiastic popular following.”

Finally, the purpose of the watchdog extends beyond simply making the management and execution of power transparent, to making known and understood the effects of that power. This includes reporting on successes as well as failures.

Journalists have an obligation to protect this watchdog freedom by not demeaning it in frivolous use or exploiting it for commercial gain.

It must provide a forum for public criticism and compromise

The news media are common carriers of public discussion, and this responsibility forms a basis for special privileges that news and information providers receive from democratic societies.

These privileges can involve subsidies for distribution or research and development (lower postal rates for print, use of public spectrum by broadcasters, development and management of the Internet) to laws protecting content and free speech (copyright, libel, and shield laws).

These privileges, however, are not pre-ordained or perpetual. Rather, they are conferred because of the need for an abundant supply of information. They are predicated on the assumption that journalism – because of its principles and practices – will supply a steady stream of higher quality content that citizens and government will use to make better decisions.

Traditionally, this covenant has been between news organizations and government. The new forms of digital media, however, place a responsibility on everyone who “publishes” content – whether for profit or for personal satisfaction – in the public domain.

The raw material cast into the marketplace of ideas sustains civic dialogue and serves society best when it consists of verified information rather than just prejudice and supposition.

Journalism should also attempt to fairly represent varied viewpoints and interests in society and to place them in context rather than highlight only the conflicting fringes of debate. Accuracy and truthfulness also require that the public discussion not neglect points of common ground or instances where problems are not just identified but also solved.

Journalism, then, is more than providing an outlet for discussion or adding one’s voice to the conversation. **Journalism carries with it a responsibility to improve the quality of debate by providing verified information and intellectual rigor.** A forum without regard for facts fails to inform and degrades rather than improves the quality and effectiveness of citizen decision-making.

It must strive to keep the significant interesting and relevant

Journalism is storytelling with a purpose. It should do more than gather an audience or catalogue the important. It must balance what readers know they want with what they cannot anticipate but need.

Writing coaches Roy Peter Clark and Chip Scanlan describe effective newswriting as the intersection of civic clarity, the information citizens need to function, and literary grace, which is the reporter's storytelling skill set. In other words, part of the journalist's responsibility is providing information in such a way people will be inclined to listen. **Journalists must thus strive to make the significant interesting and relevant.**

Quality is measured both by how much a work engages its audience and enlightens it. This means journalists must continually ask what information has the most value to citizens and in what form people are most likely to assimilate it. While journalism should reach beyond such topics as government and public safety, journalism overwhelmed by trivia and false significance trivializes civic dialogue and ultimately public policy.

It must keep the news comprehensive and proportional

Journalism is our modern cartography. It creates a map for citizens to navigate society.

As with any map, its value depends on a completeness and proportionality in which the significant is given greater visibility than the trivial.

Keeping news in proportion is a cornerstone of truthfulness. Inflating events for sensation, neglecting others, stereotyping, or being disproportionately negative all make a less reliable map. **The most comprehensive maps include all affected communities, not just those with attractive demographics.** The most complete stories take into account diverse backgrounds and perspectives.

Though proportion and comprehensiveness are subjective, their ambiguity does not lessen their significance.

Its practitioners must be allowed to exercise their personal conscience

Doing journalism, whether as a professional writing for a news organization or as an online contributor in the public space, involves **one's moral compass and demands a personal sense of ethics and responsibility.**

Because “news” is important, those who provide news have a responsibility to voice their personal conscience out loud and allow others to do so as well. They must be willing to question their own work and to differ with the work of others if fairness and accuracy demand they do so.

News organizations do well to nurture this independence by encouraging individuals to speak their minds. Conversation and debate stimulate the intellectual diversity of minds and voices necessary to understand and accurately cover an increasingly diverse society. Having a diverse newsroom does little if those different voices are not spoken or heard.

It's also a matter of self-interest. Employees encouraged to raise their hands may “save the boss from himself” or protect the news organization's reputation by pointing out errors, flagging important omissions, questioning misguided assumptions, or even revealing wrongdoing.

Having a sense of ethics is perhaps most important for the individual journalist or online contributor.

Increasingly, those who produce “the news” work in isolation, whether from a newsroom cubicle, the scene of a story, or their home office. They may file directly to the public without the safety net of editing, a second set of eyes, or the collaboration of others. While crowdsourcing by the audience may catch and correct errors or misinformation, the reputation of the author and the quality of public dialogue are nevertheless damaged.

Citizens, who shape news production by the choices they make, have rights when it comes to news, but they also have responsibilities — even more so as they become producers and editors themselves.

The average person now, more than ever, works like a journalist.

Writing a blog entry, commenting on a social media site, sending a tweet, or “liking” a picture or post, likely involves a shorthand version of the journalistic process. One comes across information, decides whether or not it’s believable, assesses its strength and weaknesses, determines if it has value to others, decides what to ignore and what to pass on, chooses the best way to share it, and then hits the “send” button.

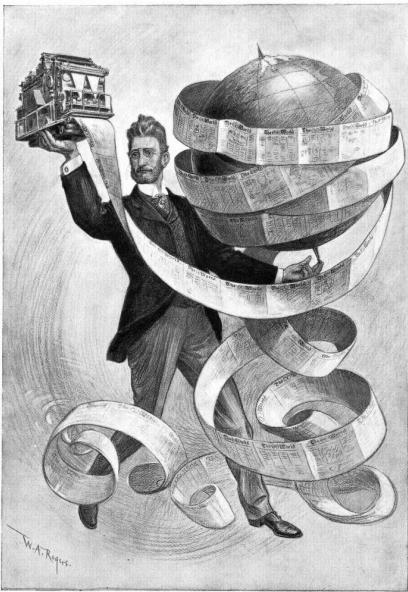
Though this process may take only a few moments, it’s essentially what reporters do.

Two things, however, separate this journalistic-like process from an end product that *is* “journalism.” The first is motive and intent. The purpose of journalism is to give people the information they need to make better decisions about their lives and society. The second difference is that journalism involves the conscious, systematic application of a discipline of verification to produce a “functional truth,” as opposed to something that is merely interesting or informative. Yet while the process is critical, it’s the end product – the “story” – by which journalism is ultimately judged.

Today, when the world is awash in information and news is available any time everywhere, a new relationship is being formed between the suppliers of journalism and the people who consume it.

The new journalist is no longer a gatekeeper who decides what the public should and should not know. The individual is now his or her own circulation manager and editor. To be relevant, journalists must now verify information the consumer already has or is likely to find and then help them make sense of what it means and how they might use it.

Thus, write Kovach and Rosenstiel, “**The first task of the new journalist/sense maker is to verify what information is reliable and then order it so people can grasp it efficiently.**” A part of this new journalistic responsibility is “to provide citizens with the tools they need to extract knowledge for themselves from the undifferentiated flood of rumor, propaganda, gossip, fact, assertion, and allegation the communications system now produces.”



AMERICAN EDITORS. II.—JOSEPH PULITZER

© 2001 HARPWEEK®

The use of data and computation in journalism is not new. Joseph Pulitzer himself wrote about how important "data" was to journalism.

NORTH AMERICAN REVIEW

No. DLXX.

MAY, 1904.

THE COLLEGE OF JOURNALISM.

A Review of Criticisms and Objections—Reflections Upon the Power, the Progress and the Prejudices of the Press—Why Specialized Concentration and Education at College Would Improve the Character and Work of Journalists and So Promote the Welfare of the Republic.

"The man who writes, the man who month in and month out, week in and week out, day in and day out, furnishes the material which is to shape the thoughts of our people, is essentially the man who more than any other determines the character of the people and the kind of government this people shall possess."

—PRESIDENT ROOSEVELT, April 7, 1904.

BY JOSEPH PULITZER.

THE editor of the NORTH AMERICAN REVIEW has asked me to reply to an article recently printed in its pages criticising the College of Journalism which it has been my pleasure to found and permanently to endow in Columbia University. In complying with his request I have enlarged the scope of the reply to include all other criticisms and misgivings, many honest, some shallow, some based on misunderstanding, but the most representing only prejudice and ignorance. If my comment upon these criticisms shall seem to be diffuse and perhaps repetitious, my apology is that—alas!—I am compelled to write by voice, not by pen, and to revise the proofs by ear, not by eye—a somewhat difficult task.

Some of my critics have called my scheme "visionary." If it

VOL. CLXXVIII.—NO. 570.

41

Copyright, 1904, by THE NORTH AMERICAN REVIEW PUBLISHING COMPANY. All Rights Reserved.

642

THE NORTH AMERICAN REVIEW.

be so I can at least plead that it is a vision I have cherished long, thought upon deeply and followed persistently. Twelve years ago I submitted the idea to President Low of Columbia, when it was declined. I have ever since continued to perfect and organize the scheme in my mind, until it is now accepted. In examining the criticisms and misgivings I have endeavored to do so without prejudice, anxious only to find the truth. I admit that the difficulties are many, but after weighing them all impartially I am more firmly convinced than ever of the ultimate success of the idea. Before the century closes schools of journalism will be generally accepted as a feature of specialized higher education, like schools of law or of medicine.

And now for our critics and objectors:

MUST A JOURNALIST BE "BORN?"

They object, the critics and cavillers, that a "newspaper man" must depend solely upon natural aptitude, or, in the common phrase, that he must be "born, not made."

Perhaps the critics can name some great editor, born full-winged like Mercury, the messenger of the gods? I know of none. The only position that occurs to me which a man in our Republic can successfully fill by the simple fact of birth is that of an idiot. Is there any other position for which a man does not demand and receive training—training at home, training in schools and colleges, training by master craftsmen, or training through bitter experience—through the burns that make the child dread the fire, through blunders costly to the aspirant?

This last is the process by which the profession of journalism at present obtains its recruits. It works by natural selection and the survival of the fittest, and its failures are strewn along the wayside.

The "born editor" who has succeeded greatly without special preparation is simply a man with unusual ability and aptitude for his chosen profession, with great power of concentration and sustained effort. He is one who loves his work and puts his whole heart and mind into it. He is in the strictest sense an educated man, but he has merely substituted *self-education* for education by others, making up for any deficiencies in his training by the unreserved sacrifice of strength, energy and pleasure. Even in his case might it not be an advantage to have a system of instruction that

In data he wrote we could find "romance, human interest, humor and fascinating revelations."

In ways that Pulitzer probably could not have imagined, over the last few decades, data and computing have become part of our everyday lives.

They regulate and shape our interactions with the physical and virtual worlds. Organizations increasingly equate (though not without problems)

"data release" with transparency. Sensing (sound, light, air quality) is cheap and plentiful, and easily deployed by the general public. Our actions online generate vast quantities of digital data.

And increasingly, computer systems exercise real power in the world through the insertion of machine learning (statistical models, artificial intelligence) alongside or in place of human decision making. In all of this, we can find new ways to ask questions about the world, how it's organized and how it functions. But the keys to this new digitized kingdom are data, code and algorithms. The curiosity, the questioning spirit, journalism students develop in their reporting classes finds an outlet in new and unexpected ways, mediated by data, code and algorithms.

Hence, a class in computational journalism. It is simply a response to our new condition of living in a computational society.

ciliation, of permanent State arbitration tribunals, of standing arbitration agreements between labor-unions and employers, and of the long line of international settlements leading up to the establishment of the world's court of arbitration at The Hague, would furnish material in themselves for a full and most valuable course of study for a journalist.

STATISTICS.

Everybody says that statistics should be taught. But how?

Statistics are not simply figures. It is said that nothing lies like figures—except facts. You want statistics to tell you the truth. You can find truth there if you know how to get at it, and romance, human interest, humor and fascinating revelations as well. The journalist must know how to find all these things—truth, of course, first. His figures must bear examination. It is much better to underestimate than to overstate his case, so that his critics and not himself may be put to confusion when they challenge him to verify his comparisons.

He must not read his statistics blindly; he must be able to test them by knowledge and by common sense. He must always be on the alert to discover how far they can actually be trusted—and what they really mean. The analysis of statistics to get at the essential truth of them has become a well-developed science, whose principles are systematically taught. And what a fascinating science it is! What romance can equal the facts of our national growth?

Statistics making an impact

John Pullinger

House of Commons Library, London, UK

[The address of the President, delivered to The Royal Statistical Society on Wednesday, June 26th, 2013]

Summary. Statistics provides a special kind of understanding that enables well-informed decisions. As citizens and consumers we are faced with an array of choices. Statistics can help us to choose well. Our statistical brains need to be nurtured: we can all learn and practise some simple rules of statistical thinking. To understand how statistics can play a bigger part in our lives today we can draw inspiration from the founders of the Royal Statistical Society. Although in today's world the information landscape is confused, there is an opportunity for statistics that is there to be seized. This calls for us to celebrate the discipline of statistics, to show confidence in our profession, to use statistics in the public interest and to champion statistical education. The Royal Statistical Society has a vital role to play.

Keywords: Chartered Statistician; Citizenship; Economic growth; Evidence; getstats; Justice; Open data; Public good; The state; Wise choices

1. Introduction

Dictionarys trace the source of the word statistics from the Latin ‘status’, the state, to the Italian ‘statista’, one skilled in statecraft, and on to the German ‘Statistik’, the science dealing with data about the condition of a state or community. The *Oxford English Dictionary* brings ‘statistics’ into English in 1787. Florence Nightingale held that ‘the thoughts and purpose of the Deity are only to be discovered by the statistical study of natural phenomena... the application of the results of such study [is] the religious duty of man’ (Pearson, 1924).

Statistics is about matters of the highest importance in human affairs. It is about our comprehension of the world around us and how that comprehension affects the lives we live: our health, wealth and wellbeing; our understanding of the state of our planet and the condition of our local communities; our choices at work and at play.

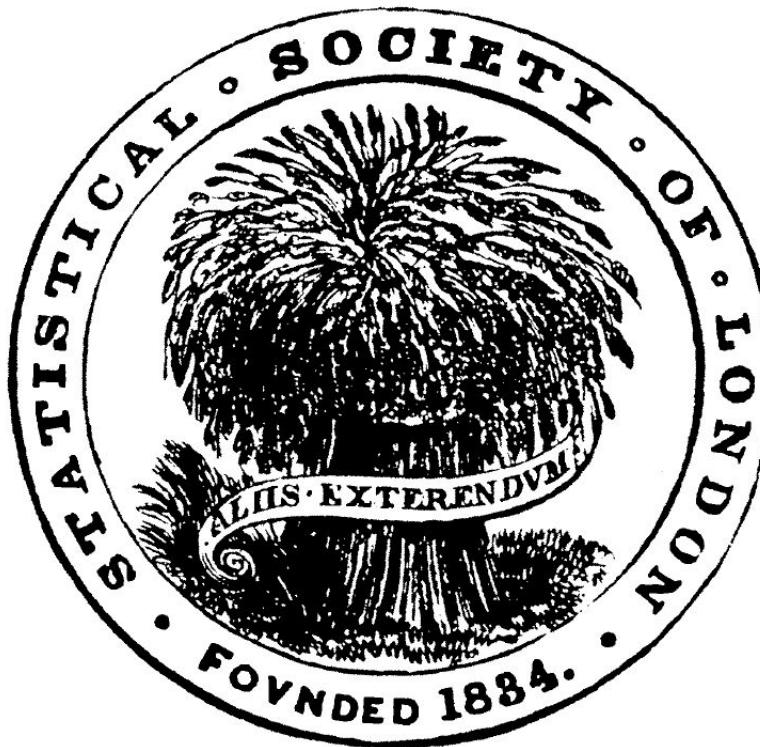
In this address I shall explore the interface between statistics and decision making. I shall identify the principles which give statistics its extraordinary ability to shape our thinking about how we view the world. I shall highlight the leading role played by the Royal Statistical Society in developing the discipline and the profession of statistics since 1834 and draw attention to why it matters so much at this time.

I shall conclude with a call to action for statistics to give people the power to make good choices. A call that chimes with the objectives of this, the International Year of Statistics (see <http://www.statistics2013.org/about-us/>), which is being promoted by the American Statistical Association, the Institute of Mathematical Statistics, the International Biometric Society and the International Statistical Institute (and the Bernoulli Society) as well as the Royal Statistical Society along with over 2000 other organizations around the world:

Address for correspondence: John Pullinger, House of Commons Library, London, SW1A 0AA, UK.
E-mail: j.pullinger@rss.org.uk

© 2013 Royal Statistical Society

0964–1998/13/176819



Principles

Anyone seems to be able to come up with a number these days. What does it really mean? Does it mean anything at all? Is it just confusing us, trying to pull the wool over our eyes?

We can all learn and practise some simple rules of statistical thinking.

Statistics are provisional. Findings can always be improved. We must be willing to recognize the validity of the figures that are available today but ready to applaud rather than criticize when tomorrow there is better information available through which to revise them.

Statistics are uncertain. We should be sceptical of those peddling impossible guarantees, rather than demanding them, and celebrate those who tell us about risk and imprecision.

Statistics aim for impartiality. In a world of spin, where numbers are used to substantiate assertions rather than illuminate reality, we must be doggedly determined to root out bias and to ask ourselves whether something is really what it seems.

Statistics help us to join the dots. We must be modest in jumping to conclusions about cause and effect but appreciate the real value that comes from finding patterns that help us to ask good questions.

Everyone can do this. But not everyone can be a statistician. We need also to appreciate that those who are statisticians abide by principles of professionalism that should give confidence in what they say. We should look carefully at the provenance of statistical claims to see whether they are coming from people or organizations that have signed up to the necessary standards of conduct and have the necessary technical training and degree of expertise.

The principles of statistics have been captured in a variety of ways. The late Roger Jowell, who was a great mentor to me and an inspiration to many, was the leading light in the development of the International Statistical Institute's code of ethics (International Statistical Institute, 2010). Within this code is a set of ethical principles which set **a high bar for professional conduct:**

Principles

- (a) **pursuing objectivity, without fear or favour and guarding against ‘predictable misinterpretation or misuse’;**
- (b) clarifying obligations and roles, statisticians should take care to **stay within their area of competence**, and to seek advice, as appropriate, from others with the relevant expertise;
- (c) **assessing alternatives** impartially, assessing the respective merits and limitations of alternatives;
- (d) avoiding financial or personal conflicts of interest, exploring the likely consequences of collecting and disseminating various types of data and results;
- (e) rejecting any attempt to establish a predetermined outcome, including contractual conditions contingent on such an outcome;

- (f) guarding privileged information, although this prohibition is not to be extended to statistical methods and procedures utilized to conduct the inquiry or to produce published data;
- (g) exhibiting professional competence, upgrading professional knowledge and skills;
- (h) maintaining confidence in statistics, **alerting potential users of the results to the limits of their reliability and applicability;**
- (i) exposing and reviewing methods and findings through independent assessment;
- (j) communicating ethical principles;
- (k) bearing responsibility for the integrity of the discipline, not deceiving or knowingly misrepresenting or attempting to prevent reporting of misconduct or obstructing the scientific or scholarly research of others;
- (l) **protecting the interests of subjects against potentially harmful effects of participation**, including keeping confidential identities and records of all subjects or respondents.

THE FUTURE OF DATA ANALYSIS¹

By JOHN W. TUKEY

Princeton University and Bell Telephone Laboratories

I. General Considerations	2
1. Introduction	2
2. Special growth areas	3
3. How can new data analysis be initiated?	4
4. Sciences, mathematics, and the arts	5
5. Dangers of optimization	7
6. Why optimization?	8
7. The absence of judgment	9
8. The reflection of judgment upon theory	10
9. Teaching data analysis	11
10. Practicing data analysis	13
11. Facing uncertainty	13
II. Spotty Data	14
12. What is it?	14
13. An appropriate step forward	15
14. Trimming and Winsorizing samples	17
15. How soon should such techniques be put into service?	19
III. Spotty Data in More Complex Situations	21
16. Modified normal plotting	21
17. Automated examination	22
18. FUNOP	22
19. FUNOR-FUNOM in a two-way table	24
20. Example of use of FUNOR-FUNOM	27
IV. Multiple-Response Data	32
21. Where are we, and why?	32
22. The case of two samples	33
23. Factor analysis: the two parts	35
24. Factor analysis: regression	36
25. Factor analysis: the middle lines	36
26. Taxonomy; classification; incomplete data	38
V. Some Other Promising Areas	39
27. Stochastic-process data	39
28. Selection and screening problems	40
29. External, internal, and confounded estimates of error	41
30. The consequences of half-normal plotting	42
31. Heterogeneous data	42
32. Two samples with unequal variability	43
VI. Flexibility of Attack	44
33. Choice of modes of expression	44
34. Sizes, nomination, budgeting	45
35. A caveat about indications	46

Received July 1, 1961.

¹ Prepared in part in connection with research sponsored by the Army Research Office through Contract DA-36-034-ORD-2207 with Princeton University. Reproduction in whole or part is permitted for any purpose of the United States Government.

between these conventional levels. To be useful, a conclusion procedure need not be precise. As working data analysts we need to recognize that this is so.

We must face up to the fact that, in any experimental science, *our certainty about what will happen in a particular situation does not usually come from directly applicable experiments or theory*, but rather comes mainly through analogy between situations which are *not known* to behave similarly. Data analysis has, of necessity, to be an experimental science, and needs therefore to adopt the attitudes of experimental science. As a consequence our choices of analytical approach will usually be guided by what is known about simpler or similar situations, rather than by what is known about the situation at hand.

Finally, we need to give up the vain hope that data analysis can be founded upon a logico-deductive system like Euclidean plane geometry (or some form of the propositional calculus) and to face up to the fact that *data analysis is intrinsically an empirical science*. Some may feel let down by this, may feel that if data analysis cannot be a logico-deductive system, it inevitably falls to the state of a crass technology. With them I cannot agree. It will still be true that there will be aspects of data analysis well called technology, but there will also be the hallmarks of stimulating science: intellectual adventure, demanding calls upon insight, and a need to find out "how things really are" by investigation and the confrontation of insights with experience.

46. How might data analysis be taught? If we carry the point of view set forth here to its logical conclusion, we would teach data analysis in a very different way from any that I know to have been tried. We would teach it like biochemistry, with emphasis on what we have learned, with some class discussion of how such things were learned perhaps, but with relegation of all question of detailed methods to the "laboratory work". If we carried through the analogy to the end, all study of detailed proofs, as well as all trials of empirical sampling or comparisons of ways of presentation would belong in "the laboratory" rather than "in class". Moreover, practice in the use of data analysis techniques would be left to other courses in which problems arose, just as applications of biochemistry are left to other courses.

It seems likely, but not certain, that this would prove to be too great a switch to consider putting into immediate effect. Even if it is too much for one step, what about taking it in two or three steps?

I can hear the war cry "cookbookery" being raised against such a proposal. If raised it would fail, because the proposal is really to go in the opposite direction from cookbookery; to teach not "what to do", nor "how we learned what to do", but rather "what we have learned". This last is at the opposite pole from "cookbookery", goes beyond "the conduct of taste-testing panels", and is concerned with "the art of cookery". Dare we adventure?

47. The impact of the computer. How vital, and how important, to the matters we have discussed is the rise of the stored-program electronic computer? In many instances the answer may surprise many by being "important but not

In data analysis we find the “...hallmarks of stimulating science: intellectual adventure, demanding calls upon insight, and **a need to find out ‘how things really are’ by investigation and the confrontation of insights with experience.**”



Subscribe to our print magazine

BECOME A MEMBER NOW »



An ode to reporter's notebooks

F

FIRST PERSON

By Philip Eil

MAY 9, 2018



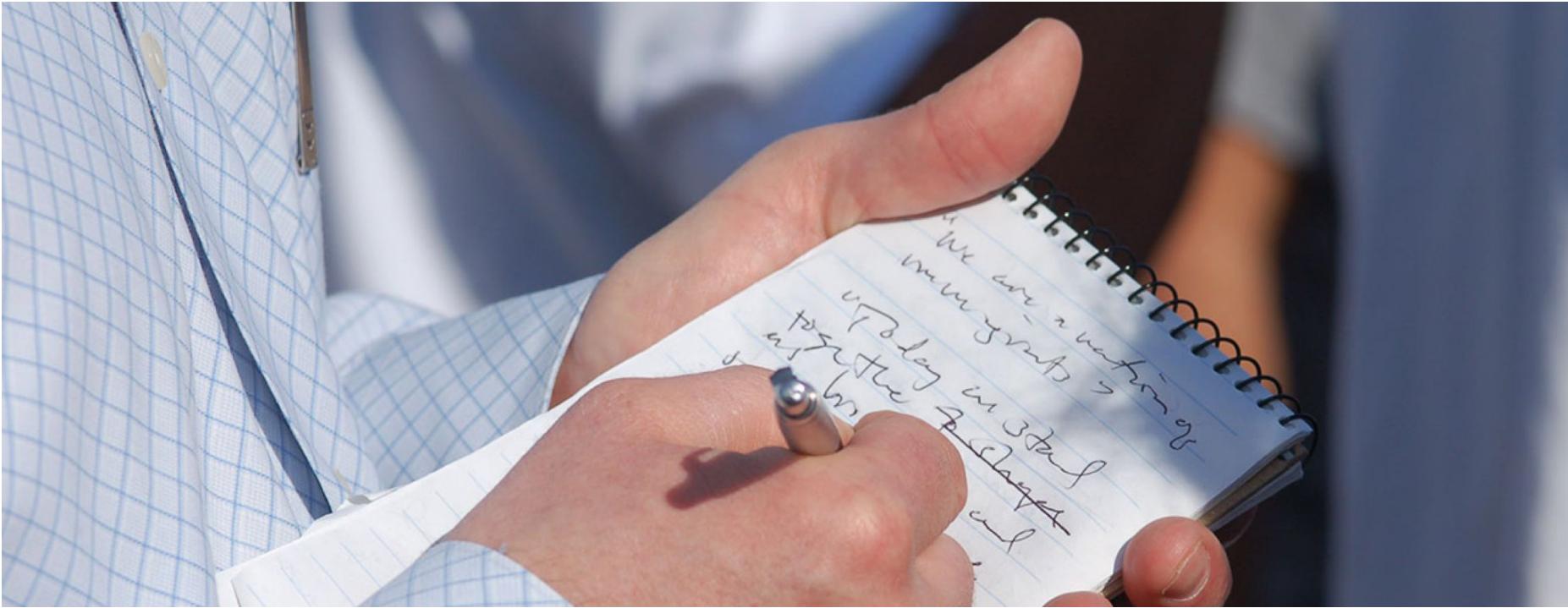
SHARE
ON TWITTER



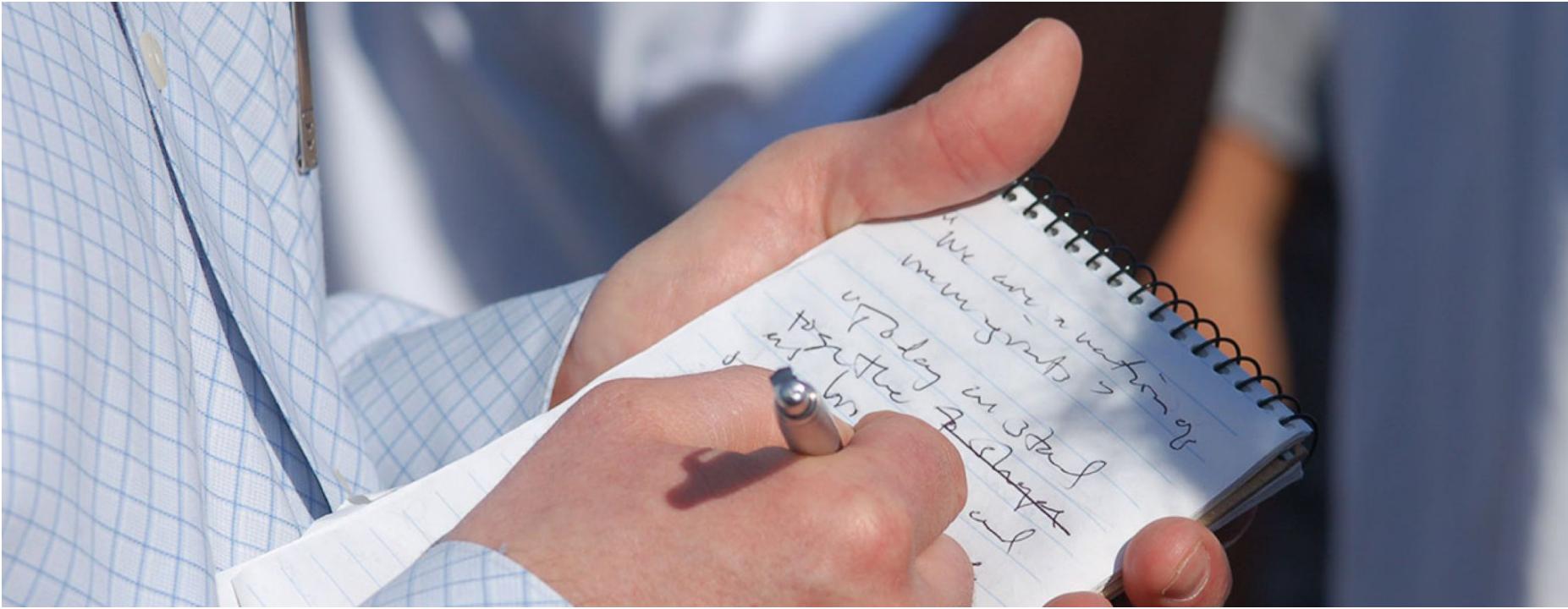
SHARE
ON FACEBOOK

TEN YEARS AGO THIS MONTH, I became a reporter.

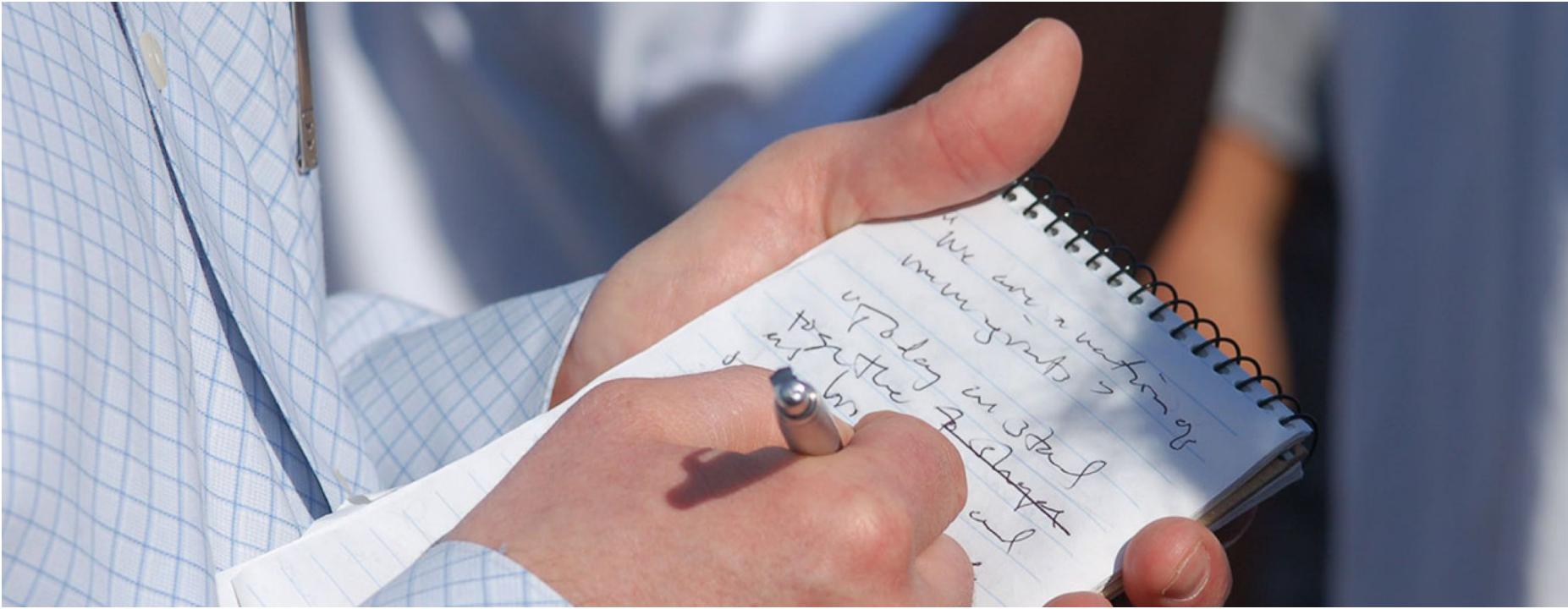
In April of 2008, I was a 23-year-old recent college grad working at a soul-stifling corporate marketing job and feeling increasingly terrified of a life filled with spreadsheets, menial tasks, khakis, and meetings with unclear purposes. Since enrolling in a night class on creative nonfiction a few months earlier, though, I had begun to see a different path. And by May, I had abandoned my cubicle, driven across the country, and started an internship at the alt-weekly, the *San Francisco Bay*



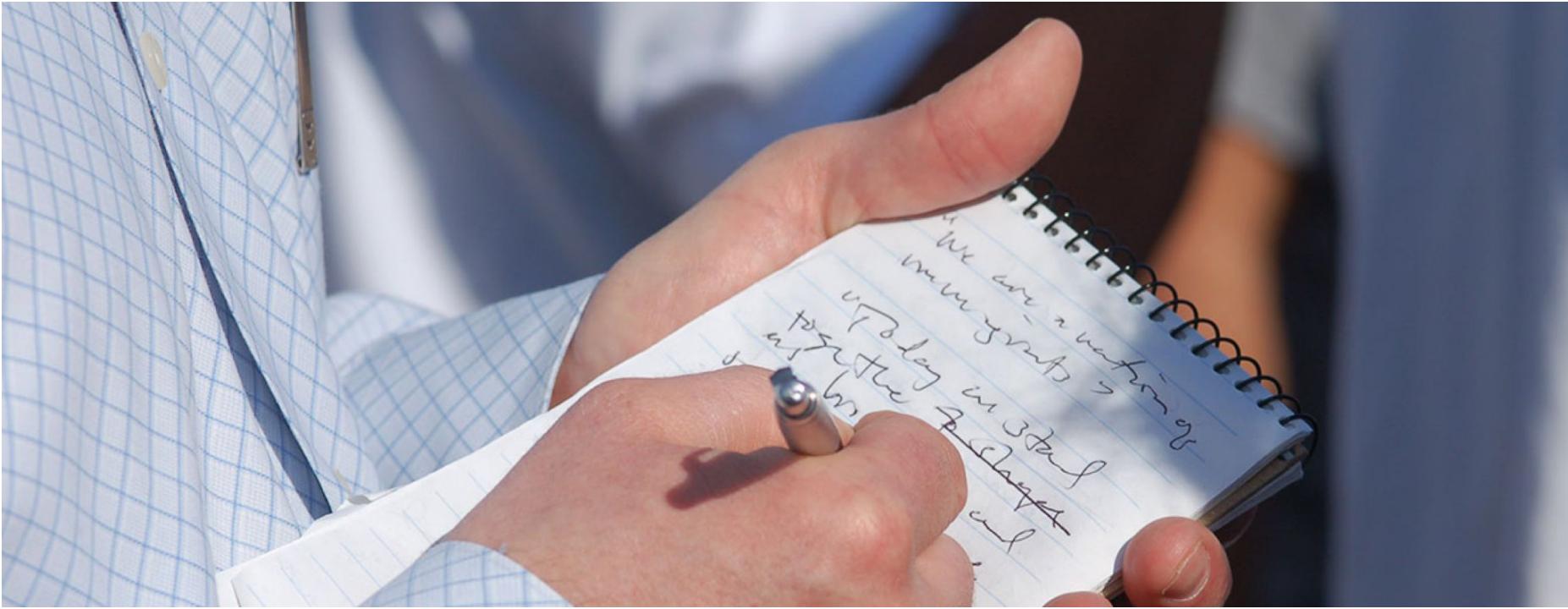
While preparing for this briefing, I reread a beautiful piece in CJR by Philip Eil, a freelancer from Providence Rhode Island. His article was titled **“An Ode to Reporter’s Notebooks.”** For Eil, a reporter’s notebook is more than paper and cardboard. It symbolizes the power and uniqueness of journalism as a profession. In his words, this object is “a license to talk to interesting people” and “a ticket to interesting places.”



"[T]he reporter's notebook is the low-tech device I use to capture the sights, smells, sounds, feelings, tastes, and other impressions of the world. **To report is to be alert and alive in a particular time and place...**"



My training is not in journalism but statistics. When I arrived at the J School, I was told over and over that journalism is less an academic discipline than it is a habit of mind. We cultivate in our students a curiosity about the world, a restless, questioning spirit that asks why things look as they do.

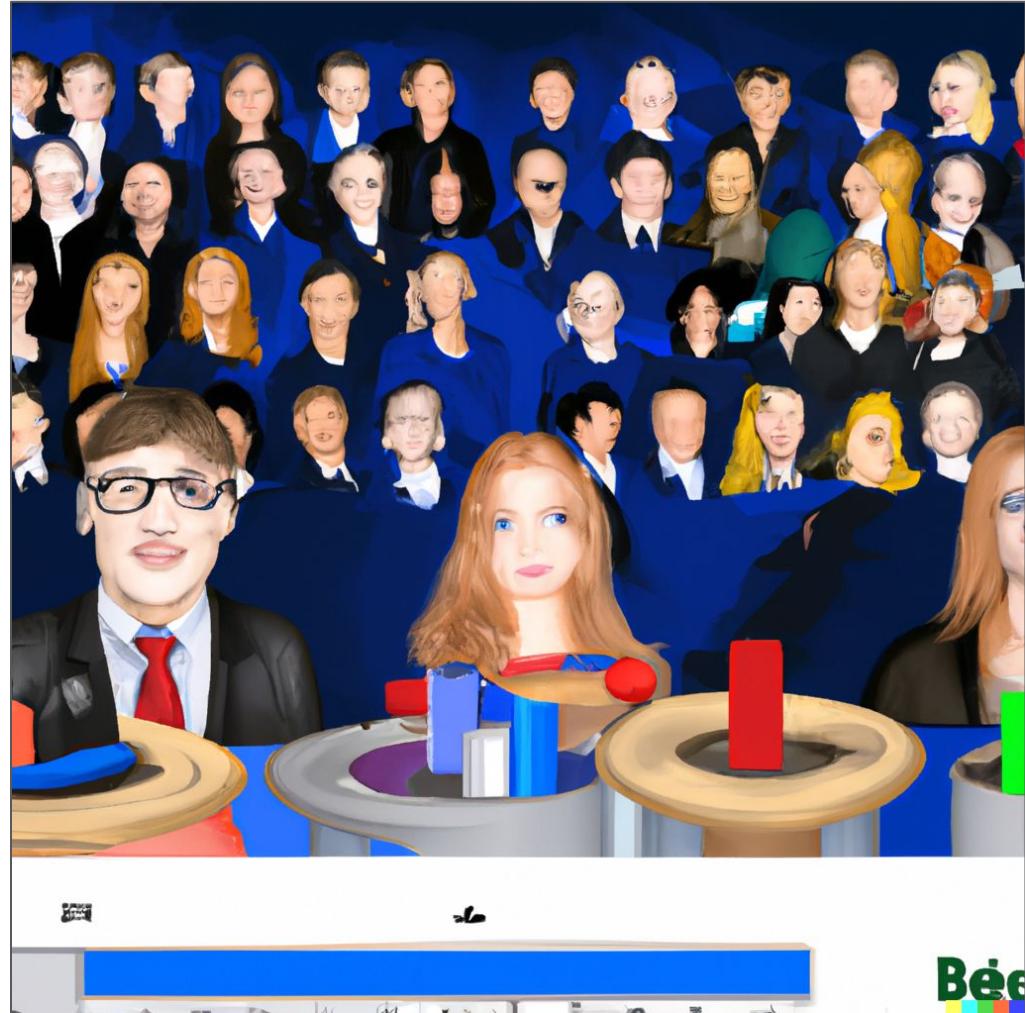


And my role as a statistician in a journalism school is to add computational lines of inquiry to this habit of mind. If you take one thing from today's briefings, it's that just as the reporter's notebook represents a call to be alert and alive in a particular time and place, for me in my teaching, **computation extends that alertness and curiosity to the virtual world of data, code and algorithms.**

JOURNALIST'S BRIEFING

EXERCISE: IN THE NEWS

Introducing some of the complexities of
Computational Journalism



AI in the News...

VICE

Conservatives Are Panicking About AI Bias, Think ChatGPT Has Gone 'Woke'

All AI systems carry biases, and ChatGPT allegedly being "woke" is far from the most dangerous one.

By Matthew Gault

January 11, 2023, 9:00am

Share | Tweet | Save

JOHN LUND / GETTY

GOLF

PRO SHOP NEWS INSTRUCTION GEAR TRAVEL & LIFESTYLE LEADERSBOARDS VIDEOS & PODCASTS

I talked golf with viral chatbot ChatGPT. The results blew my mind

BY ALAN RASTALO | JANUARY 18, 2023

hi there, are you a golfer?

No, I am not a golfer.

would you like to learn how to play golf?

Yes, I would like to learn how to play golf.

what's stopping you?

I don't have access to a golf course or golf equipment. Additionally, I need knowledge

difficult golf course in the world?

The most difficult golf course in the world is widely considered to be the Par 3 Cape Kidnappers in New Zealand. This course features narrow fairways, deep bunkers, deep ravines, and long carry distances. The wind also plays a major factor, making the course even more challenging.

interesting, I wasn't aware cape kidnappers even had a par 3 course

Yes, Cape Kidnappers is home to a par 3 course which has been dubbed as the most difficult golf course in the world. Many golfers have attempted the course and have

Bloomberg

• Live Now Markets Economics Industries Technology Politics Wealth Pursuits Opinion Businessweek Equality Green

Technology + Work Shift

ChatGPT Will Be the Calculator for Writing, Top Economist Says

■ Stanford professor says AI could reduce mindless, rote work
■ Predicts new technology will "augment our ability to write"

The OpenAI website ChatGPT. Photographer: Gabby Jones/Bloomberg

By Philip Aldrick

January 18, 2023 at 11:26 AM EST

Axios Miami

Latest stories Membership Jobs

Miami real estate agents experiment with ChatGPT

Martin Vassolo

f t in e

OpenAI

ChatGPT: Optimizing Language Models for Dialogue

OpenAI is a company created by DALL-E and GPT-3 creators. It is a research organization focused on developing AI systems that can understand and generate human-like language. The company was founded in 2015 and is based in San Francisco, California. It has since expanded to include offices in London, Berlin, Paris, and Tokyo. OpenAI's mission is to develop safe and useful AI systems that benefit humanity as a whole. The company's work includes research into generative models, reinforcement learning, and other areas of AI development.

DAVOS, SWITZERLAND

WILL.I.AM: I think ChatGPT is awesome. LIVES THROUGH EDUCATION

yahoo/finance

2:01 / 6:44

WILL.I.AM: I think ChatGPT is awesome. LIVES THROUGH EDUCATION

WORLD ECONOMIC FORUM Davos

VQ

Davos 2023: ChatGPT 'a great co-pilot for creatives,' will.i.am says

Alexandra Garfinkle · Senior Reporter

Thu, January 19, 2023 at 8:25 AM EST · 2 min read

What about for journalism?

The Washington Post
Democracy Dies in Darkness

A news site used AI to write articles. It was a journalistic disaster.

The tech site CNET sent a chill through the media world when it tapped artificial intelligence to produce surprisingly lucid news stories. But now its human staff is writing a lot of corrections.

 By Paul Farhi

January 17, 2023 at 6:03 p.m. EST



Turns out the robots aren't better than us at journalism — yet. (iStock)

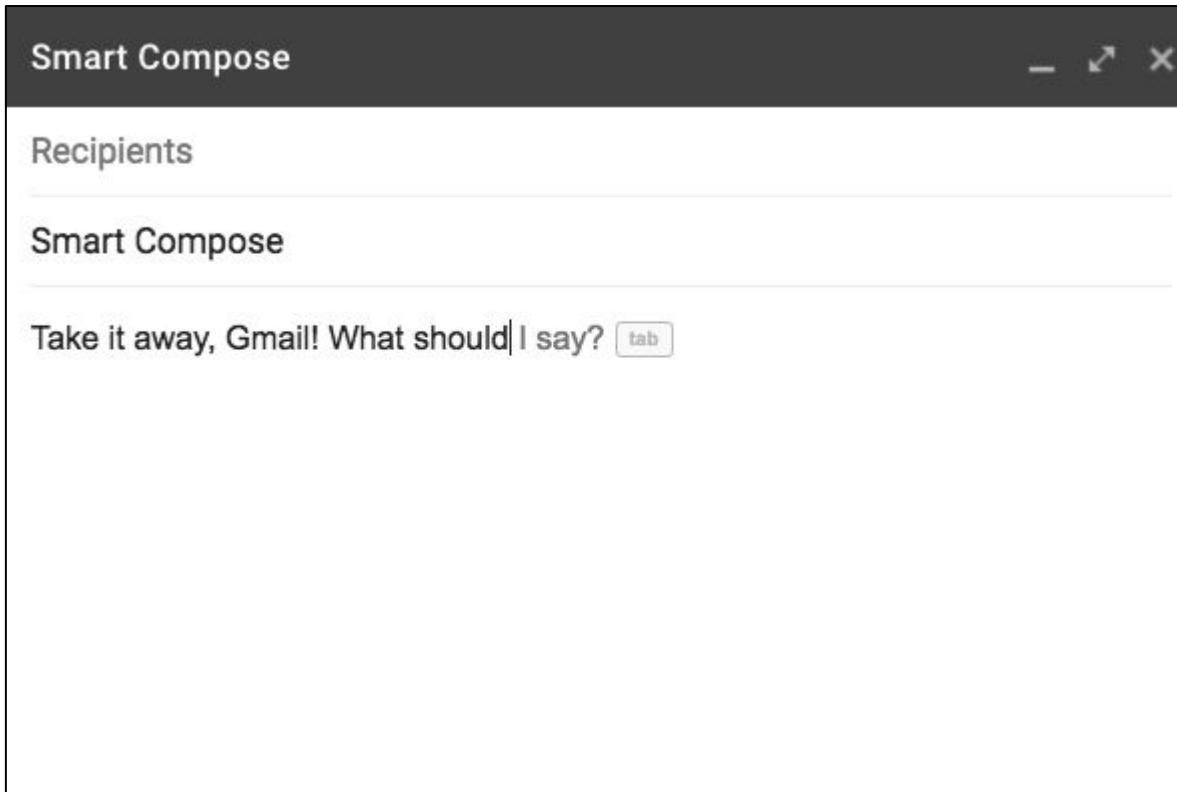
Language Models

Trained to predict the next word in a sentence:

The cat is chasing the _____



Does it have any real applications?



Playground - OpenAI API x +

beta.openai.com/playground

Overview Documentation Examples Playground

Help Columbia University

Playground Load a preset... Save View code Share ...

In April of 2008, I was a 23-year-old recent college grad working at a soul-stifling corporate marketing job and feeling increasingly terrified of a life filled with spreadsheets, menial tasks, khakis, and meetings with unclear purposes. Since enrolling in a night class on creative nonfiction a few months earlier, though, I had begun to see a different path.

That April, I walked out of my corporate office for the last time, with a plan to take a trip that would change my life. I was headed to the island of Maui with a one-way ticket, a new laptop, and a dream of becoming a writer.

The weeks that followed were full of exploration, adventure, and learning. I hiked to the top of dormant volcanoes, swam with sea turtles, camped on the beach, and went on a whale-watching tour. I also began to write, tapping out stories and essays on my laptop while watching the sun set into the Pacific Ocean.

The trip was a turning point in my life. I realized that no matter how scared I was, I had the courage to make a bold move, and that I had the capacity to make a living doing something I loved. I returned home with a new sense of purpose. I began to freelance and build my writing career, eventually leaving my corporate job behind for good.

Eight years later, I'm still writing and traveling. I've published essays in national magazines, written for top websites, and taught writing workshops. I'm also a published author and I have my own travel blog. I'm

⚠ Warning: Your text ends in a trailing space, which causes worse performance due to how the API splits text into tokens.

Submit ⌂ ⌂ ⌂ ⌂ ⌂ 335 Inject start text

Mode: List, Grid, Card

Model: text-davinci-003

Temperature: 0.7

Maximum length: 256

Stop sequences: Enter sequence and press Tab

Top P: 1

Frequency penalty: 0

Presence penalty: 0

Best of: 1

Language Models are Few-Shot Learners

Tom B. Brown*	Benjamin Mann*	Nick Ryder*	Melanie Subbiah*
Jared Kaplan [†]	Prafulla Dhariwal	Arvind Neelakantan	Pranav Shyam
Amanda Askell	Sandhini Agarwal	Ariel Herbert-Voss	Gretchen Krueger
Rewon Child	Aditya Ramesh	Daniel M. Ziegler	Jeffrey Wu
Christopher Hesse	Mark Chen	Eric Sigler	Mateusz Litwin
Benjamin Chess	Jack Clark	Christopher Berner	Scott Gray
Sam McCandlish	Alec Radford	Ilya Sutskever	Dario Amodei

OpenAI

Abstract

Recent work has demonstrated substantial gains on many NLP tasks and benchmarks by pre-training on a large corpus of text followed by fine-tuning on a specific task. While typically task-agnostic in architecture, this method still requires task-specific fine-tuning datasets of thousands or tens of thousands of examples. By contrast, humans can generally perform a new language task from only a few examples or from simple instructions – something which current NLP systems still largely struggle to do. Here we show that scaling up language models greatly improves task-agnostic, few-shot performance, sometimes even reaching competitiveness with prior state-of-the-art fine-tuning approaches. Specifically, we train GPT-3, an autoregressive language model with 175 billion parameters, 10x more than any previous non-sparse language model, and test its performance in the few-shot setting. For all tasks, GPT-3 is applied without any gradient updates or fine-tuning, with tasks and few-shot demonstrations specified purely via text interaction with the model. GPT-3 achieves strong performance on many NLP datasets, including translation, question-answering, and cloze tasks, as well as several tasks that require on-the-fly reasoning or domain adaptation, such as unscrambling words using a novel word in a sentence, or performing 3-digit arithmetic. At the same time, GPT-3 performs worse on datasets where GPT-3’s few-shot learning still struggles, as well as datasets where GPT-3 faces methodological issues related to training on large web corpora. Finally, we find that GPT-3 can generate samples of news articles which human evaluators have difficulty distinguishing from articles written by humans. We discuss broader societal impacts of this finding and of GPT-3 in general.

*Equal contribution

[†]Johns Hopkins University, OpenAI

Author contributions listed at end of paper.

Language Models are Unsupervised Multitask Learners

Alec Radford *[†] Jeffrey Wu *[†] Rewon Child[†] David Luan[†] Dario Amodei **[†] Ilya Sutskever **[†]

Abstract

Natural language processing tasks, such as question answering, machine translation, reading comprehension, and summarization, are typically approached with supervised learning of task-specific datasets. We demonstrate that language models begin to learn these tasks without any explicit supervision when trained on a new dataset of millions of webpages called WebText. When conditioned on a document plus questions, the answers generated by the language model reach 55 F1 on the CoQA dataset - matching or exceeding the performance of 3 out of 4 baseline systems without using the 127,000+ training examples. The capacity of the language model is essential to the success of zero-shot task transfer and increasing it improves performance in a log-linear fashion across tasks. Our largest model, GPT-2, is a 1.5B parameter Transformer that achieves state of the art results on 7 out of 8 tested language modeling datasets in a zero-shot setting but still underfits WebText. Samples from the model reflect these improvements and contain coherent paragraphs of text. These findings suggest a promising path towards building language processing systems which learn to perform tasks from their naturally occurring demonstrations.

1. Introduction

Machine learning systems now excel (in expectation) at tasks they are trained for by using a combination of large datasets, high-capacity models, and supervised learning (Krizhevsky et al., 2012) (Sutskever et al., 2014) (Amodei et al., 2016). Yet these systems are brittle and sensitive to slight changes in the data distribution (Recht et al., 2018) and task specification (Kirkpatrick et al., 2017). Current systems are better characterized as narrow experts rather than

*[†]Equal contribution [†]OpenAI, San Francisco, California, United States. Correspondence to: Alec Radford <alec@openai.com>.

competent generalists. We would like to move towards more general systems which can perform many tasks – eventually without the need to manually create and label a training dataset for each one.

The dominant approach to creating ML systems is to collect a dataset of training examples demonstrating correct behavior for a desired task, train a system to imitate these behaviors, and then test its performance on independent and identically distributed (IID) held-out examples. This has served well to make progress on narrow experts. But the often erratic behavior of captioning models (Lake et al., 2017), reading comprehension systems (Jia & Liang, 2017), and image classifiers (Alcorn et al., 2018) on the diversity and variety of possible inputs highlights some of the shortcomings of this approach.

Our suspicion is that the prevalence of single task training on single domain datasets is a major contributor to the lack of generalization observed in current systems. Progress towards robust systems with current architectures is likely to require training and measuring performance on a wide range of domains and tasks. Recently, several benchmarks have been proposed such as GLUE (Wang et al., 2018) and decaNLP (McCann et al., 2018) to begin studying this.

Multitask learning (Caruana, 1997) is a promising framework for improving general performance. However, multitask training in NLP is still nascent. Recent work reports modest performance improvements (Yogatama et al., 2019) and the two most ambitious efforts to date have trained on a total of 10 and 17 (dataset, objective) pairs respectively (McCann et al., 2018) (Bowman et al., 2018). From a meta-learning perspective, each (dataset, objective) pair is a single training example sampled from the distribution of datasets and objectives. Current ML systems need hundreds to thousands of examples to induce functions which generalize well. This suggests that multitask training many need just as many effective training pairs to realize its promise with current approaches. It will be very difficult to continue to scale the creation of datasets and the design of objectives to the degree that may be required to brute force our way there with current techniques. This motivates exploring additional setups for performing multitask learning.

The current best performing systems on language tasks

Zero-shot

The model predicts the answer given only a natural language description of the task. No gradient updates are performed.



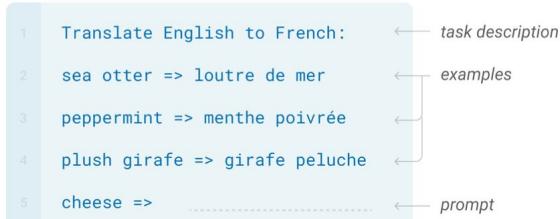
One-shot

In addition to the task description, the model sees a single example of the task. No gradient updates are performed.



Few-shot

In addition to the task description, the model sees a few examples of the task. No gradient updates are performed.



“Having a huge, massively pre-trained and generalist model that knows and has encapsulated a lot of information is the real key to the paradigm shift!

In fact, all this allows you to interact with the model in a very similar way to how you would ask a **human being** to perform a particular task, that is, giving instructions and at most a few examples, all in natural language! The ability to perform tasks with a couple of examples (versus the thousands or millions needed for deep learning systems) is called **few shot learning** and is the revolution that has emerged from the training of Large Language Models.

How to give instructions to the model? Simply by typing in a **prompt** what I want the model to do.

In this way you can do a lot of things, from writing stories to solving math problems. All this without the need to retrain the model! Here the main difficulty becomes writing prompts in an optimal and optimized way for the task.”

An example — New York State Department of Health Mpox Reporting

The screenshot shows a web browser window with the title "Monkeypox" and the URL "https://www.health.ny.gov/diseases/communicable/zoonoses/monkeypox/". The browser interface includes a navigation bar with back, forward, and search buttons, and a toolbar with various icons. A watermark for "Wayback Machine" is visible on the left, indicating the page was captured on July 14, 2022, at 14:22 UTC. The main content area has a purple header with links for "Department of Health", "Individuals/Families", "Providers/Professionals", "Health Facilities", "Health Data", and "Search". Below the header, a breadcrumb trail shows "Home Page > Zoonoses > Monkeypox". The main section features a large heading "Monkeypox" and a sub-section "Update: Monkeypox in New York State – July 2022". A callout box contains the text: "As of July 14 2022, a total of 414 confirmed orthopoxvirus/monkeypox cases - a designation established by the Centers for Disease Control and Prevention (CDC) - have been identified with 389 in New York City, 12 in Westchester County, 4 in Suffolk County, 3 in Nassau County, 1 in Sullivan County, 1 in Chemung County, 1 in Rockland County, 1 in Erie County, 1 in St. Lawrence County, and 1 case in Monroe County." The text below discusses monkeypox as a rare viral infection and provides information on who is at risk and how it spreads.

Monkeypox

Update: Monkeypox in New York State – July 2022

As of July 14 2022, a total of 414 confirmed orthopoxvirus/monkeypox cases - a designation established by the Centers for Disease Control and Prevention (CDC) - have been identified with 389 in New York City, 12 in Westchester County, 4 in Suffolk County, 3 in Nassau County, 1 in Sullivan County, 1 in Chemung County, 1 in Rockland County, 1 in Erie County, 1 in St. Lawrence County, and 1 case in Monroe County.

Monkeypox is a rare, viral infection that does not usually cause serious illness. However, it can result in hospitalization or death. That's why health officials in New York, the U.S., and around the world are monitoring cases of monkeypox in areas that do not usually report monkeypox infections, including in New York State.

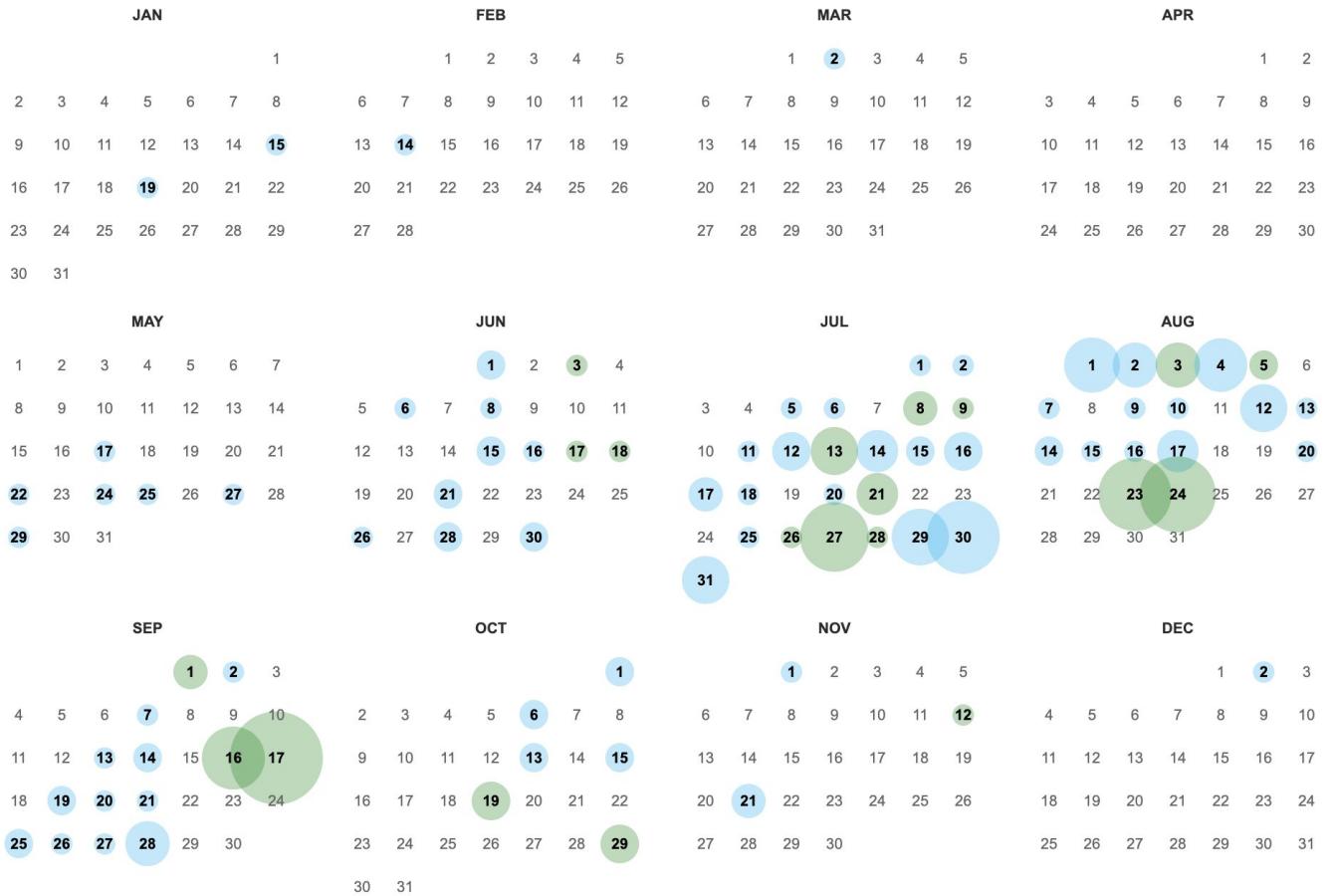
While New Yorkers should not be alarmed, everyone should stay informed about monkeypox. This means understanding the symptoms, how it spreads, and what to do if you are exposed.

Who is at risk for contracting monkeypox?

Monkeypox spreads through close, physical contact between people. This means anyone can get monkeypox. However, based on the current outbreak, certain populations are being affected by monkeypox more than others, including men who have sex with men (MSM).

Based on previous outbreaks of monkeypox around the world, some groups may also be at heightened risk for severe outcomes if they contract monkeypox. This includes people with weakened immune systems, elderly New Yorkers, young children under 8 years of age, and pregnant people.

What are the symptoms of monkeypox?	How does monkeypox spread?	How can I protect myself?
Symptoms of monkeypox can include:	Monkeypox is spread through close, physical	New Yorkers can protect themselves by taking



INTERNET ARCHIVE
WayBack Machine



Mpox

CDC > Poxvirus > Mpox > 2022 Outbreak Cases & Data

Mpox

Your Health

2022 Outbreak Cases & Data

2022 U.S. Map & Case Count

U.S. Case Trends

2022 Global Map & Case Count

Outbreak Reproduction Number Estimates

Vaccine Administration & Effectiveness

Laboratory Testing

Case Demographics & Symptoms

TPOXX Patient Data

Behavioral Survey

Technical Report

Past U.S. Cases & Outbreaks

Healthcare Professionals

Laboratories

Health Departments

Veterinarians

Communication Resources



CDC is updating webpages with the term "mpox" to reduce stigma and other issues associated with prior terminology. This change is aligned with the recent [World Health Organization](#) decision.

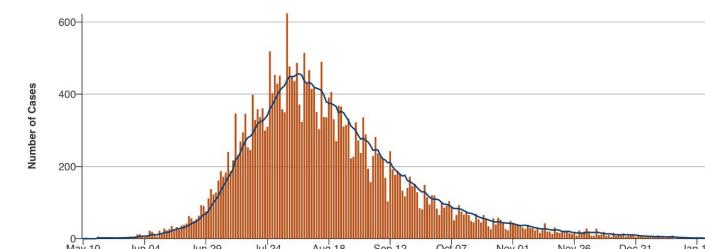
U.S. Mpox Case Trends Reported to CDC

Data as Reported to CDC as of 18 Jan 2023 2:00 PM EDT

Español | Print

Trends of mpox cases reported to CDC during the 2022 outbreak by date*

Daily Mpox Cases and 7 Day Daily Average



Data Table



Download Data (CSV)

Data are updated every Wednesday as soon as they are reviewed and verified. Historical cases may be added as additional information is obtained through case investigations.

* Date is defined as the earliest date available regarding the case, beginning with date of illness onset, diagnosis date, positive laboratory test report date, CDC call center reporting date, or case data entry date into CDC's emergency response common operating platform, DCIPHER.

Monkeypox

INTERNET ARCHIVE Wayback Machine https://www.health.ny.gov/diseases/communicable/zoonoses/monkeypox/ Go JUN JUL AUG 14 2021 2022 2023 About this capture

440 captures 6 Oct 2010 - 2 Dec 2022

Department of Health Individuals/Families Providers/Professionals Health Facilities Health Data Search

You are Here: [Home Page](#) > [Zoonoses](#) > Monkeypox

Monkeypox

Update: Monkeypox in New York State – July 2022

As of July 14 2022, a total of 414 confirmed orthopoxvirus/monkeypox cases - a designation established by the Centers for Disease Control and Prevention (CDC) - have been identified with 389 in New York City, 12 in Westchester County, 4 in Suffolk County, 3 in Nassau County, 1 in Sullivan County, 1 in Chemung County, 1 in Rockland County, 1 in Erie County, 1 in St. Lawrence County, and 1 case in Monroe County.

Monkeypox is a rare, viral infection that does not usually cause serious illness. However, it can result in hospitalization or death. That's why health officials in New York, the U.S., and around the world are monitoring cases of monkeypox in areas that do not usually report monkeypox infections, including in New York State.

While New Yorkers should not be alarmed, everyone should stay informed about monkeypox. This means understanding the symptoms, how it spreads, and what to do if you are exposed.

Who is at risk for contracting monkeypox?

Monkeypox spreads through close, physical contact between people. This means anyone can get monkeypox. However, based on the current outbreak, certain populations are being affected by monkeypox more than others, including men who have sex with men (MSM).

Based on previous outbreaks of monkeypox around the world, some groups may also be at heightened risk for severe outcomes if they contract monkeypox. This includes people with weakened immune systems, elderly New Yorkers, young children under 8 years of age, and pregnant people.

What are the symptoms of monkeypox?	How does monkeypox spread?	How can I protect myself?
Symptoms of monkeypox can include:	Monkeypox is spread through close, physical	New Yorkers can protect themselves by taking

Here is a paragraph summarizing the latest monkeypox statistics for New York state:

As of July 13 2022, a total of 359 confirmed orthopoxvirus/monkeypox cases - a designation established by the Centers for Disease Control and Prevention (CDC) - have been identified with 336 in New York City, 12 in Westchester County, 4 in Suffolk County, 2 in Nassau County, 1 in Sullivan County, 1 in Chemung County, 1 in Rockland County, 1 in Erie County, and 1 in St. Lawrence County.

Create a CSV where each row contains information on a places, recording both the place and the case count.

Place, Case Count

The screenshot shows the OpenAI Playground interface on a Mac OS X desktop. The window title is "Playground - OpenAI API". The URL in the address bar is "beta.openai.com/playground". The main content area displays two prompts and their generated outputs. The first prompt is: "Here is a paragraph summarizing the latest monkeypox statistics for New York state:". The output is: "As of July 13 2022, a total of 359 confirmed orthopoxvirus/monkeypox cases - a designation established by the Centers for Disease Control and Prevention (CDC) - have been identified with 336 in New York City, 12 in Westchester County, 4 in Suffolk County, 2 in Nassau County, 1 in Sullivan County, 1 in Chemung County, 1 in Rockland County, 1 in Erie County, and 1 in St. Lawrence County." The second prompt is: "Create a CSV where each row contains information on a places, recording both the place and the case count." The output is: "Place, Case Count". The right side of the interface features various configuration options for the AI model, including Mode (set to text-davinci-003), Temperature (0.7), Maximum length (256), Stop sequences (empty), Top P (1), Frequency penalty (0), Presence penalty (0), and Best of (1). A status message at the bottom left says "Looking for ChatGPT? Try it now". The character count of the current input is shown as 140.

Here are three paragraphs summarizing the latest monkeypox statistics for New York state:

1. As of July 11 2022, a total of 238 confirmed orthopoxvirus/monkeypox cases - a designation established by the Centers for Disease Control and Prevention (CDC) - have been identified with 223 in New York City, 7 in Westchester County, 4 in Suffolk County, 1 in Nassau County, 1 in Sullivan County, 1 in Chemung County, and 1 in Rockland County.
2. As of July 13 2022, a total of 359 confirmed orthopoxvirus/monkeypox cases - a designation established by the Centers for Disease Control and Prevention (CDC) - have been identified with 336 in New York City, 12 in Westchester County, 4 in Suffolk County, 2 in Nassau County, 1 in Sullivan County, 1 in Chemung County, 1 in Rockland County, 1 in Erie County, and 1 in St. Lawrence County.
3. As of July 14 2022, a total of 414 confirmed orthopoxvirus/monkeypox cases - a designation established by the Centers for Disease Control and Prevention (CDC) - have been identified with 389 in New York City, 12 in Westchester County, 4 in Suffolk County, 3 in Nassau County, 1 in Sullivan County, 1 in Chemung County, 1 in Rockland County, 1 in Erie County, 1 in St. Lawrence County, and 1 case in Monroe County.

Where are the biggest changes in monkeypox cases occurring?

The screenshot shows a web browser window with the title "CSV of Monkeypox Cases". The main content area displays the following text:

Westchester County, 4 in Suffolk County, 2 in Nassau County, 1 in Sullivan County, 1 in Chemung County, 1 in Rockland County, 1 in Erie County, and 1 in St. Lawrence County.

3. As of July 14 2022, a total of 414 confirmed orthopoxvirus/monkeypox cases - a designation established by the Centers for Disease Control and Prevention (CDC) - have been identified with 389 in New York City, 12 in Westchester County, 4 in Suffolk County, 3 in Nassau County, 1 in Sullivan County, 1 in Chemung County, 1 in Rockland County, 1 in Erie County, 1 in St. Lawrence County, and 1 case in Monroe County.

Where are the biggest changes in monkeypox cases occurring?

The AI response below states:

The biggest changes in monkeypox cases appear to be occurring in New York City, as the case count has increased from 223 on July 11th to 389 on July 14th, representing a increase of 166 cases in 3 days. Additionally, there have been increases in the case counts in Westchester County, Nassau County, and Erie County. New cases have also been identified in St. Lawrence County and Monroe County during this time period.

At the bottom of the screen, a footer bar includes links for "Clear conversations", "Dark mode", "OpenAI Discord", "Updates & FAQ", and "Log out".

Monkeypox

INTERNET ARCHIVE Wayback Machine https://www.health.ny.gov/diseases/communicable/zoonoses/monkeypox/ Go JUN JUL AUG 14 2021 2022 2023 About this capture

440 captures 6 Oct 2010 - 2 Dec 2022

Department of Health Individuals/Families Providers/Professionals Health Facilities Health Data Search

You are Here: [Home Page](#) > [Zoonoses](#) > Monkeypox

Monkeypox

Update: Monkeypox in New York State – July 2022

As of July 14 2022, a total of 414 confirmed orthopoxvirus/monkeypox cases - a designation established by the Centers for Disease Control and Prevention (CDC) - have been identified with 389 in New York City, 12 in Westchester County, 4 in Suffolk County, 3 in Nassau County, 1 in Sullivan County, 1 in Chemung County, 1 in Rockland County, 1 in Erie County, 1 in St. Lawrence County, and 1 case in Monroe County.

Monkeypox is a rare, viral infection that does not usually cause serious illness. However, it can result in hospitalization or death. That's why health officials in New York, the U.S., and around the world are monitoring cases of monkeypox in areas that do not usually report monkeypox infections, including in New York State.

While New Yorkers should not be alarmed, everyone should stay informed about monkeypox. This means understanding the symptoms, how it spreads, and what to do if you are exposed.

Who is at risk for contracting monkeypox?

Monkeypox spreads through close, physical contact between people. This means anyone can get monkeypox. However, based on the current outbreak, certain populations are being affected by monkeypox more than others, including men who have sex with men (MSM).

Based on previous outbreaks of monkeypox around the world, some groups may also be at heightened risk for severe outcomes if they contract monkeypox. This includes people with weakened immune systems, elderly New Yorkers, young children under 8 years of age, and pregnant people.

What are the symptoms of monkeypox?	How does monkeypox spread?	How can I protect myself?
Symptoms of monkeypox can include:	Monkeypox is spread through close, physical	New Yorkers can protect themselves by taking

Monkeypox

INTERNET ARCHIVE Wayback Machine https://www.health.ny.gov/diseases/communicable/zoonoses/monkeypox/ Go JUN JUL SEP 29 2021 2022 2023 About this capture

474 captures 6 Oct 2010 - 2 Dec 2022

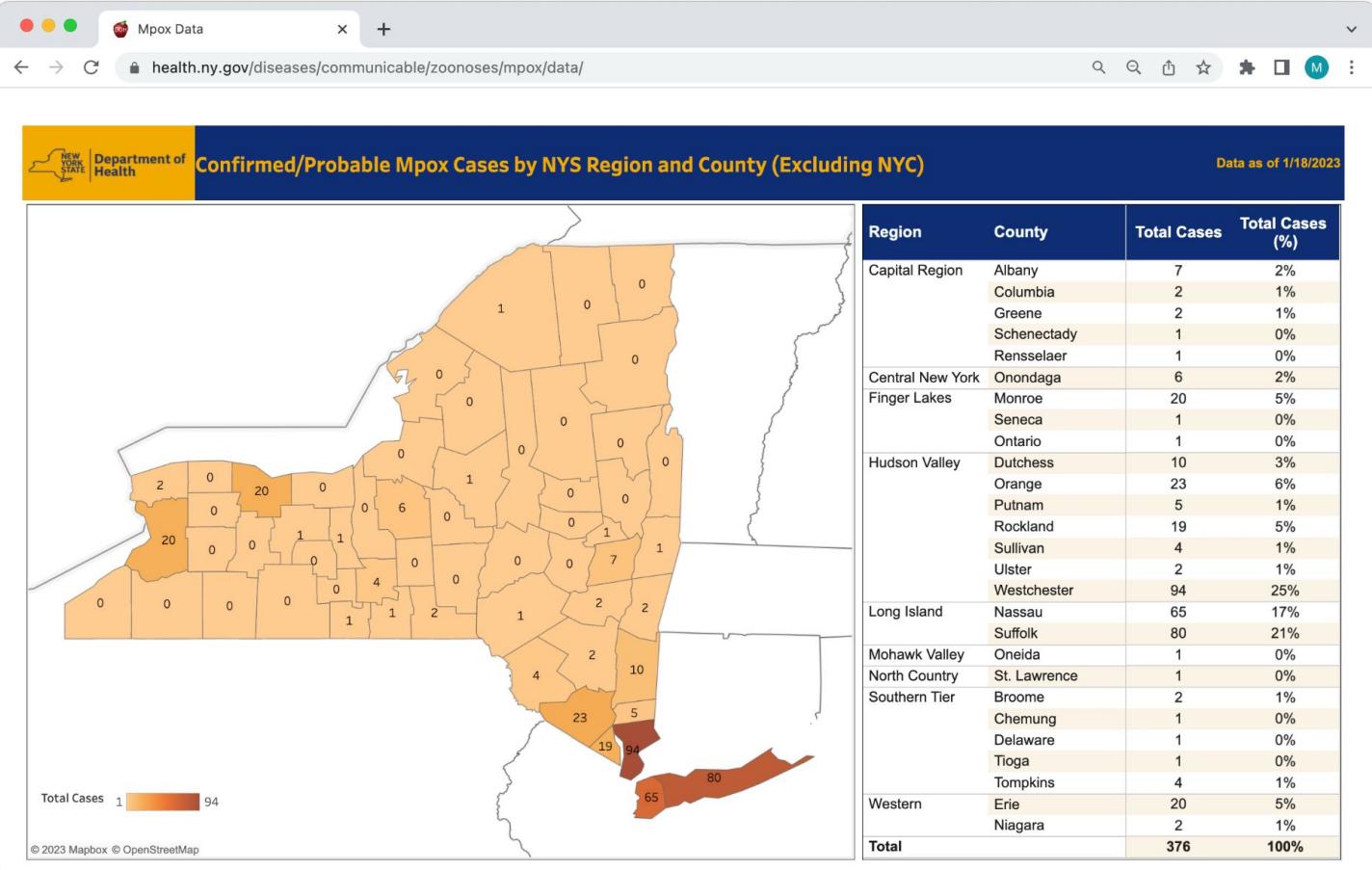
What to Know Vaccine Information Provider Information Events and Videos

Update: Monkeypox in New York State – July 2022

As of July 29 2022, a total of 1,383 confirmed orthopoxvirus/monkeypox cases - a designation established by the Centers for Disease Control and Prevention (CDC).

 Click here for full county listing

County of Residence	Cases
Albany	1
Broome	1
Chemung	1
Columbia	1
Dutchess	3
Erie	4
Greene	1
Monroe	6
Nassau	8
Niagara	1
Orange	3
Rockland	4
Waiting for we... nce	1





2022 Weekly Summary of Monkeypox Cases in New York State (Excluding NYC) Residents

The data include probable and confirmed monkeypox cases in New York State, not including NYC residents. The data are inclusive of case reports received by the NYSDOH through 12/21/2022 at 7:00 AM, eastern time, and is subject to change as follow up activities are ongoing.

Additional information on daily case counts in New York State can be found at:
<https://health.ny.gov/diseases/communicable/zoonoses/monkeypox/data/>

Confirmed/Probable Monkeypox Cases, in NYS (Excluding NYC)

	N	(%)
Total	373	100%
Age Group		
Under 18 years	3	1%
18-24 years	50	13%
25-29 years	77	21%
30-34 years	91	24%
35-39 years	45	12%
40-44 years	39	10%
45-49 years	24	6%
50-54 years	23	6%
55 years and above	21	6%
Gender^{1, 2}		
Male	354	95%
Female	12	3%
TGNCNB	3	1%
Unknown	4	1%
Sexual Orientation¹		
Straight	43	12%
Lesbian or gay	161	43%
Bisexual	33	9%
Other	3	1%
Unknown	133	36%
Race and Ethnicity¹		
Hispanic/Latino	119	32%
White, non-Hispanic	104	28%
Black/African American, non-Hispanic	70	19%
American Indian or Alaskan Native	1	0%
Asian or Pacific Islander	8	2%
Other	16	4%
Unknown	55	15%

¹ Unknown due to pending investigation or non-disclosure during investigation.

² Includes the following gender categories: male, female, transgender, gender nonconforming and nonbinary (TGNCNB). People whose current gender identity differs from their sex assigned at birth are considered TGNCNB. Sex and gender information are collected from self-reports, diagnosing provider or electronic lab reporting. This information may or may not accurately reflect self-identification.

Playground - OpenAI API

beta.openai.com/playground

Overview Documentation Examples Playground

Help Columbia University

Playground

Load a preset... Save View code Share ...

Maximum length 797

Stop sequences Enter sequence and press Tab

Top P 1

Frequency penalty 0

Presence penalty 0

Best of 1

Inject start text

Inject restart text

Show probabilities

Here is a paragraph summarizing the latest monkeypox statistics for New York state:
As of July 13 2022, a total of 359 confirmed orthopoxvirus/monkeypox cases - a designation established by the Centers for Disease Control and Prevention (CDC) - have been identified with 336 in New York City, 12 in Westchester County, 4 in Suffolk County, 2 in Nassau County, 1 in Sullivan County, 1 in Chemung County, 1 in Rockland County, 1 in Erie County, and 1 in St. Lawrence County.

Create a CSV where each row contains information on a place, recording bot the place and the case count.
Place, Case Count

Submit

140 Off Show probabilities

Playground - OpenAI API

beta.openai.com/playground

Overview Documentation Examples Playground

Help Columbia University

Playground

Load a preset... Save View code Share ...

Maximum length 797

Stop sequences Enter sequence and press Tab

Top P 1

Frequency penalty 0

Presence penalty 0

Best of 1

Inject start text

Inject restart text

Show probabilities

Here is a paragraph summarizing the latest monkeypox statistics for New York state:
As of July 13 2022, a total of 359 confirmed orthopoxvirus/monkeypox cases - a designation established by the Centers for Disease Control and Prevention (CDC) - have been identified with 336 in New York City, 12 in Westchester County, 4 in Suffolk County, 2 in Nassau County, 1 in Sullivan County, 1 in Chemung County, 1 in Rockland County, 1 in Erie County, and 1 in St. Lawrence County.

Create a CSV where each row contains information on a place, recording bot the place and the case count.
Place, Case Count
New York City, 336
Westchester County, 12
Suffolk County, 4
Nassau County, 2
Sullivan County, 1
Chemung County, 1
Rockland County, 1
Erie County, 1
St. Lawrence County, 1

Submit

140 Off Show probabilities

Playground - OpenAI API

beta.openai.com/playground

Overview Documentation Examples Playground Help Columbia University

Playground

Load a preset... Save View code Share ...

Maximum length 797

Stop sequences Enter sequence and press Tab

Top P 1

Frequency penalty 0

Presence penalty 0

Best of 1

Inject start text

Inject restart text

Show probabilities Off

Here are three paragraphs summarizing the latest monkeypox statistics for New York state:

1. As of July 11 2022, a total of 238 confirmed orthopoxvirus/monkeypox cases - a designation established by the Centers for Disease Control and Prevention (CDC) - have been identified with 223 in New York City, 7 in Westchester County, 4 in Suffolk County, 1 in Nassau County, 1 in Sullivan County, 1 in Chemung County, and 1 in Rockland County.

2. As of July 11 2022, a total of 238 confirmed orthopoxvirus/monkeypox cases - a designation established by the Centers for Disease Control and Prevention (CDC) - have been identified with 223 in New York City, 7 in Westchester County, 4 in Suffolk County, 1 in Nassau County, 1 in Sullivan County, 1 in Chemung County, and 1 in Rockland County.

3. As of July 13 2022, a total of 359 confirmed orthopoxvirus/monkeypox cases - a designation established by the Centers for Disease Control and Prevention (CDC) - have been identified with 336 in New York City, 12 in Westchester County, 4 in Suffolk County, 2 in Nassau County, 1 in Sullivan County, 1 in Chemung County, 1 in Rockland County, 1 in Erie County, and 1 in St. Lawrence County.

Where are the biggest changes in monkeypox cases occurring?

Submit

Playground - OpenAI API

beta.openai.com/playground

Overview Documentation Examples Playground

Load a preset... Save View code Share ...

Maximum length 797

Stop sequences Enter sequence and press Tab

Top P 1

Frequency penalty 0

Presence penalty 0

Best of 1

Inject start text

Inject restart text

Show probabilities Off

Here are three paragraphs summarizing the latest monkeypox statistics for New York state:

1. As of July 11 2022, a total of 238 confirmed orthopoxvirus/monkeypox cases - a designation established by the Centers for Disease Control and Prevention (CDC) - have been identified with 223 in New York City, 7 in Westchester County, 4 in Suffolk County, 1 in Nassau County, 1 in Sullivan County, 1 in Chemung County, and 1 in Rockland County.

2. As of July 11 2022, a total of 238 confirmed orthopoxvirus/monkeypox cases - a designation established by the Centers for Disease Control and Prevention (CDC) - have been identified with 223 in New York City, 7 in Westchester County, 4 in Suffolk County, 1 in Nassau County, 1 in Sullivan County, 1 in Chemung County, and 1 in Rockland County.

3. As of July 13 2022, a total of 359 confirmed orthopoxvirus/monkeypox cases - a designation established by the Centers for Disease Control and Prevention (CDC) - have been identified with 336 in New York City, 12 in Westchester County, 4 in Suffolk County, 2 in Nassau County, 1 in Sullivan County, 1 in Chemung County, 1 in Rockland County, 1 in Erie County, and 1 in St. Lawrence County.

Where are the biggest changes in monkeypox cases occurring?

The majority of monkeypox cases are occurring in New York City, with the number of cases increasing from 223 to 336 from July 11-13. There are also small increases in the number of cases in Westchester, Suffolk, and Nassau counties.

Unlike most previous AI, that needed to be trained to do a specific task with lots of training data, GPT is not trained to do anything in particular, but it can do pretty well on a diverse set of tasks:

Summarization

Translation

Deriving keywords

Generating ideas

Rewriting text

... and so much more

Summarization

Mayor Adams Releases Preliminary Budget for Fiscal Year 2024

January 12, 2023

Video available at: https://youtu.be/3zh0R_x_sY



FY24 Preliminary Budget Includes Investments in Affordable Housing, Street Cleanliness, and Public Safety

Budget Underscores Mayor's Commitment to Strong Fiscal Management to Ensure Robust and Equitable Recovery

NEW YORK – New York City Mayor Eric Adams today released New York City's balanced \$102.7 billion **Preliminary Budget for Fiscal Year 2024 (FY24)**. The budget reflects the mayor's ongoing commitment to promoting an equitable recovery by making investments in affordable housing, keeping city streets clean, ensuring the safety of communities, and promoting a greener, healthier city. The Preliminary Budget also doubles down on Mayor Adams' commitment to fiscal responsibility by spending limited city resources wisely amid the ongoing economic and fiscal challenges facing the city, state, and nation, and maintaining the city's budget reserves at a record level of \$8.3 billion.

"As our city continues its recovery, our administration continues to make investments in our core priorities — including public safety, affordable housing, and clean streets — while exercising strong fiscal management," said **Mayor Adams**. "By asking agencies to self-fund new needs with preexisting resources, the Fiscal Year 2024 Preliminary Budget continues our strong track record of making prudent use of taxpayer dollars while continuing to 'Get Stuff Done' for New Yorkers."

Playground - OpenAI API

beta.openai.com/playground

Overview Documentation Examples Playground

Help Columbia University

Playground

Load a preset... Save View code Share ...

Mode

Model text-davinci-003

Temperature 0.7

Maximum length 256

Stop sequences Enter sequence and press Tab

Top P 1

Frequency penalty 0

Presence penalty 0

Best of 1

Summarize the following press release from Mayor Adams of New York City:

NEW YORK – New York City Mayor Eric Adams today released New York City's balanced \$102.7 billion Preliminary Budget for Fiscal Year 2024 (FY24). The budget reflects the mayor's ongoing commitment to promoting an equitable recovery by making investments in affordable housing, keeping city streets clean, ensuring the safety of communities, and promoting a greener, healthier city. The Preliminary Budget also doubles down on Mayor Adams' commitment to fiscal responsibility by spending limited city resources wisely amid the ongoing economic and fiscal challenges facing the city, state, and nation, and maintaining the city's budget reserves at a record level of \$8.3 billion.

"As our city continues its recovery, our administration continues to make investments in our core priorities — including public safety, affordable housing, and clean streets — while exercising strong fiscal management," said Mayor Adams. "By asking agencies to self-fund new needs with preexisting resources, the Fiscal Year 2024 Preliminary Budget continues our strong track record of making prudent use of taxpayer dollars while continuing to 'Get Stuff Done' for New Yorkers."

The Preliminary Budget was crafted in response to slowing economic growth, which impacts revenue, and fiscal uncertainties. The city's revenue forecast was updated to reflect an additional \$1.7 billion in FY23 and \$738 million in FY24. The windfall increase in FY23 is driven by continued momentum from the record Wall Street activity in 2021, though the city expects that continued slowing growth in the economy will lead to slowing growth in tax revenue over the financial plan.

Looking for ChatGPT? Try now

Submit

1,835

<https://www.nyc.gov/office-of-the-mayor/news/028-23/mayor-adams-releases-preliminary-budget-fiscal-year-2024#0>

Playground - OpenAI API X +

beta.openai.com/playground

Overview Documentation Examples Playground Help Columbia University

Playground Load a preset... Save View code Share ...

reductions in agency operations. (\$1.62 million).
FY24 capital investments include:

Accelerating projects to meet carbon reduction targets set forward in Local Law 97 (\$259 million);
Funding high-priority street reconstruction projects related to Vision Zero, the Streets Plan, and more (\$228 million);
Transforming the Willets Point community and creating jobs as part of a public-private partnership that builds 2,500 affordable homes and creates a soccer stadium, 40,000 square feet of public open space, and a hotel (\$153 million);
Reconstructing and upgrading the Riverside Park Soldiers' and Sailors' Monument and plaza (\$62.3 million);
Installing new locks and security measures in New York City Department of Education school facilities to make them safer for students and staff (\$47.5 million);
Doubling signal installation to 200 annually and making related upgrades (\$77 million); and
Making repairs and upgrades to marine infrastructure in Staten Island and Manhattan, including piers and barges (\$46 million).

Summary: New York City Mayor Eric Adams has released the city's balanced \$102.7 billion Preliminary Budget for Fiscal Year 2024 (FY24) which makes investments in affordable housing, keeping city streets clean, ensuring the safety of communities, and promoting a greener, healthier city. The budget also doubles down on Mayor Adams' commitment to fiscal responsibility by spending limited city resources wisely amid the ongoing economic and fiscal challenges facing the city, state, and nation, while maintaining the city's budget reserves at a record level of \$8.3 billion. The budget includes investments to support universal dyslexia screenings in public schools, provide job opportunities for youth, expand enforcement against tenant harassment, and build 500,000 new homes. It also includes funding for carbon reduction targets, street reconstruction projects, public-private partnerships, signal installation, and repairs and upgrades to marine infrastructure.

Looking for ChatGPT? Try it now ↗

Submit ↻ ↶ ↷ ↶ ↷ 1,835

Mode ☰ ⌄ ☰

Model text-davinci-003

Temperature 0.7

Maximum length 256

Stop sequences Enter sequence and press Tab

Top P 1

Frequency penalty 0

Presence penalty 0

Best of 1

Translation

Translate this into spanish:

Today's announcement is the first major milestone in New York City's Offshore Wind Vision Plan. The agreement between NYCEDC, SSBMT, and Equinor will help establish New York as a leader in offshore wind and help the New York City meet its nation-leading climate goals of 100 percent clean electricity by 2040. As part of the deal, the South Brooklyn Marine Terminal will be upgraded and transformed into an operations and maintenance base, with heavy lift platforms being built on the 39th Street Pier for wind turbine staging and installation. The port will serve as a hub to support the Empire Wind and Beacon Wind offshore wind farms. Equinor has also committed to establishing a \$5 million ecosystem fund to bring more New York City residents into offshore wind careers, propel offshore wind innovation, and support a just transition.

Hoy se anunció el primer hito importante en el Plan de Visión de Energía Eólica Offshore de la Ciudad de Nueva York. El acuerdo entre NYCEDC, SSBMT y Equinor ayudará a establecer a Nueva York como líder en energía eólica offshore y ayudará a la ciudad de Nueva York a cumplir sus objetivos climáticos líderes en la nación de 100 por ciento de energía limpia para 2040. Como parte del acuerdo, el South Brooklyn Marine Terminal será mejorado y transformado en una base de operaciones y mantenimiento, con plataformas de elevación pesada siendo construidas en el muelle de la calle 39 para la instalación de turbinas eólicas. El puerto servirá como un centro para apoyar las centrales eólicas offshore Empire Wind y Beacon Wind.

Deriving Keywords

Give me 10 keywords for this news article:

Today's announcement is the first major milestone in New York City's Offshore Wind Vision Plan. The agreement between NYCEDC, SSBMT, and Equinor will help establish New York as a leader in offshore wind and help the New York City meet its nation-leading climate goals of 100 percent clean electricity by 2040. As part of the deal, the South Brooklyn Marine Terminal will be upgraded and transformed into an operations and maintenance base, with heavy lift platforms being built on the 39th Street Pier for wind turbine staging and installation. The port will serve as a hub to support the Empire Wind and Beacon Wind offshore wind farms. Equinor has also committed to establishing a \$5 million ecosystem fund to bring more New York City residents into offshore wind careers, propel offshore wind innovation, and support a just transition.

New York City, offshore wind, climate goals, clean electricity, South Brooklyn Marine Terminal, Equinor, Empire Wind, Beacon Wind, ecosystem fund, just transition

Generating Ideas

What are 3 potential controversies about the follow press release:

Today's announcement is the first major milestone in New York City's Offshore Wind Vision Plan. The agreement between NYCEDC, SSBMT, and Equinor will help establish New York as a leader in offshore wind and help the New York City meet its nation-leading climate goals of 100 percent clean electricity by 2040. As part of the deal, the South Brooklyn Marine Terminal will be upgraded and transformed into an operations and maintenance base, with heavy lift platforms being built on the 39th Street Pier for wind turbine staging and installation. The port will serve as a hub to support the Empire Wind and Beacon Wind offshore wind farms. Equinor has also committed to establishing a \$5 million ecosystem fund to bring more New York City residents into offshore wind careers, propel offshore wind innovation, and support a just transition.

- 1) The agreement between NYCEDC, SSBMT, and Equinor could be seen as controversial because it gives Equinor too much power and control over NYC's offshore wind.
- 2) Some people may see the \$5 million ecosystem fund as a way for Equinor to buy its way into the city and its offshore wind goals.
- 3) The transformation of the South Brooklyn Marine Terminal could be controversial because it will displace some businesses that are currently there.

Rewriting text for different audiences

Describe what a prion is.

Prions are infectious, self-replicating protein particles that are responsible for serious neurodegenerative diseases in humans and animals. Unlike viruses, prions do not contain nucleic acid and are not alive.

Describe in a concise manner what a prion is.

A prion is a small protein that is infectious and causes serious neurodegenerative diseases in humans and animals.

Scientifically describe what a prion is



A prion is a type of protein that can cause infectious diseases in animals and humans. Prions are abnormal, infectious proteins that are found in the brain tissue of animals and humans who have certain neurological diseases. These proteins are thought to cause the disease by causing other proteins in the brain to change shape and function abnormally.

Describe to a 5-year old what a prion is.

A prion is a type of protein that can cause serious neurological problems in humans and animals. Prions are found in the brain and are responsible for causing diseases like mad cow disease.

Extracting Data from Text

Below is a description of a podcast, "America First with Sebastian Gorka".

sebastian spends the entire first hour with legendary actor robert davi and breitbart editor-in-chief alex marlow, the distributors behind the upcoming movie "my son hunter," about the true story behind the movie, how the right can begin to retake the culture, and more

Here is a numbered list of the names of the guests on the podcast:

1. Robert Davi
2. Alex Marlow

Extract Data from Data (!)

Below are 3 lines of a CSV about environmental abuses in Brazil.

data_infracao	estado	municipio	cpf_ou_cnpj	nome_autuado	valor_multa	status_debito	sancoes_aplicadas	SG_UF
DS_CARGO	NM_CANDIDATO	NM_URNA_CANDIDATO	NR_CPF_CANDIDATO	DS_SITUACAO_CANDIDATURA				
DS_DETALHE_SITUACAO_CAND	SG_PARTIDO	DT_NASCIMENTO	NR_IDADE_DATA_POSSE	DS_GENERO	DS_GRAU_INSTRUCAO			
DS_ESTADO_CIVIL	DS_COR_RACA	DS_OCUPACAO	VR_DESPESA_MAX_CAMPANHA					
2005-08-03	CEARA FORTIM	32297009372	RAIMUNDO ANSELMO LIMA MORORÓ	50000.0	Rescisão de parcelamento extra judicial	70 64 Lei, 9605/98, 512º Decreto, 3179/1999	CE DEPUTADO FEDERAL RAIMUNDO ANSELMO LIMA MORORÓ	ANSELMO MORORÓ
							APTO DEFERIDO	AVANTE
							04/02/1960	62.0 MASCULINO SUPERIOR COMPLETO
								CASADO(A) PARDA EMPRESÁRIO
								3176572.53
2005-04-28	DISTRITO FEDERAL BRASILIA	51630427187	AURENI BRUZ DE ARAUJO	500.0	Ajuizado	11 2º - Decreto, 3179/1999,		
							70 72 Lei, 9605/98 DF DEPUTADO DISTRITAL	AURENI BRAZ VILANOVA DE SOUSA NENA BRAZ
							51630427187	APTO DEFERIDO PRTB 16/01/1972 50.0 FEMININO ENSINO MÉDIO COMPLETO
								CASADO(A) BRANCA OUTROS 1270629.01
2005-08-12	GOIAS CALDAS NOVAS	52141314100	EVANDO MAGAL ABADIA CORREIA SILVA	5000.0	Quitado por Parcelamento (Leis 12249, 12865, 12973, 12996)	70 38/60 Lei, 9605/98, 44 2º Decreto, 3179/1999, 2º B 3º Lei, 4771/65 GO DEPUTADO FEDERAL EVANDO MAGAL ABADIA CORREIA SILVA	MAGAL 52141314100	APTO DEFERIDO PATRIOTA 13/12/1970 52.0 MASCULINO ENSINO MÉDIO COMPLETO
								CASADO(A) BRANCA OUTROS 3176572.53

The names of the 3 people involved are:

RAIMUNDO ANSELMO LIMA MORORÓ
AURENI BRAZ VILANOVA DE SOUSA
EVANDO MAGAL ABADIA CORREIA SILVA

You can make metaphors, pull quotes from an interview, and so much more...

Examples

Explore what's possible with some example applications

 All categories

 **Q&A**
Answer questions based on existing knowle...

 **Grammar correction**
Corrects sentences into standard English.

 **Summarize for a 2nd grader**
Translates difficult text into simpler concep...

 **Natural language to OpenAI API**
Create code to call to the OpenAI API usin...

 **Text to command**
Translate text into programmatic commands.

 **English to other languages**
Translates English text into French, Spanish...

 **Natural language to Stripe API**
Create code to call the Stripe API using nat...

 **SQL translate**
Translate natural language to SQL queries.

 **Parse unstructured data**
Create tables from long form text

 **Classification**
Classify items into categories via example.

 **Python to natural language**
Explain a piece of Python code in human un...

 **Movie to Emoji**
Convert movie titles into emoji.

 **Calculate Time Complexity**
Find the time complexity of a function.

 **Translate programming languages**
Translate from one programming language ...

 **Advanced tweet classifier**
Advanced sentiment detection for a piece o...

 **Explain code**
Explain a complicated piece of code.

 **Keywords**
Extract keywords from a block of text.

 **Factual answering**
Guide the model towards factual answering ...

 **Ad from product description**
Turn a product description into ad copy.

 **Product name generator**
Create product names from examples word...

 **TL;DR summarization**
Summarize text by adding a 'tl;dr:' to the en...

 **Python bug fixer**
Find and fix bugs in source code.

 **Spreadsheet creator**
Create spreadsheets of various kinds of dat...

 **JavaScript helper chatbot**
Message-style bot that answers JavaScript ...

 **ML/AI language model tutor**
Bot that answers questions about language...

 **Science fiction book list maker**
Create a list of items for a given topic.

 **Tweet classifier**
Basic sentiment detection for a piece of text.

 **Airport code extractor**
Extract airport codes from text.

 **SQL request**
Create simple SQL queries.

 **Extract contact information**
Extract contact information from a block of ...

 **JavaScript to Python**
Convert simple JavaScript expressions into ...

 **Friend chat**
Emulate a text message conversation.

GPT's Limitations — Yes there are many!

- Its data includes information through about August 2020, and so it is not good at answering questions about current events.
- It doesn't know true from false (it doesn't "know" anything) and, as we saw with the CNET example, it will happily make up reasonable sounding stories — it will not recreate the kind of information finding you might perform — it won't Google something to get the right number and it.
- The data come with heavy biases — it is the web, after all — and can be offensive. It can also be put to "evil" ends.

Exercise: Given all this, what can you do with this? Work in pairs and see what you make of it. Keep track of the kinds of things you are inputting. *This is essentially coding.* It is a good exercise as you will often be called on to both report **with** new technology as well as **on** that same new technology.

Some Hints: GPT is, at its core, a random walk so you can try generating again!

What are 10 important Chinese phrases to learn and provide a translation?

1. 你好 (nǐ hǎo) - Hello
2. 我 (wǒ) - I / me
3. 你 (nǐ) - You
4. 是 (shì) - Is / am / are
5. 不 (bù) - No
6. 我们 (wǒ men) - We / us
7. 他们 (tā men) - They
8. 这 (zhè) - This
9. 那 (nà) - That
10. 了 (le) - (completes a sentence)

What are 10 important Chinese phrases to learn and provide a translation?

- 你好 - Hello
再见 - Goodbye
请问 - Excuse me
谢谢 - Thank you
没关系 - You're welcome
不客气 - You're welcome (more formal)
对不起 - Sorry
没关系 - It's okay
不用谢 - Don't mention it
别客气 - Don't be polite

Some Hints: Specify the number of list items you want

What are important cultural experiences when visiting France?

Some important cultural experiences when visiting France include visiting famous landmarks such as the Eiffel Tower and the Louvre, trying traditional French foods such as escargot and crème brûlée, and experiencing the French art and culture scene by visiting museums and galleries.

What are 10 important cultural experiences when visting France?

1. Touring the Palace of Versailles
2. Seeing the Mona Lisa at the Louvre
3. Taking a romantic stroll along the Seine River
4. Climbing to the top of the Eiffel Tower
5. Enjoying a glass of wine at a Parisian café
6. people-watching at the Champs-Elysées
7. Admiring the Notre Dame Cathedral
8. Shopping at the Galeries Lafayette
9. Taking a scenic drive through the French countryside
10. sampling French cuisine

Some Hints: Rephrase the prompt to be more precise — small changes can make a difference

Give me 5 ideas for essays to write.

1. A personal essay about your unique experiences and perspective
2. An essay discussing the impact of a particular event or person in your life
3. A creative essay exploring a particular theme or topic
4. An analytical essay breaking down and examining a particular issue or concept
5. A reflective essay sharing your thoughts and feelings on a particular topic

Give me 5 ideas for college entrance essays to write that show I am creative and resiliant.

1. How I overcame a challenge in my life
2. How I pursued a creative passion despite obstacles
3. How I found resilience in the face of adversity
4. How I turned a negative experience into a positive one
5. How I turned a difficult situation into an opportunity for growth

Some Hints: “Prompt Chaining”

What's an example of a time where someone turned a difficult situation into an opportunity for growth?



An example of someone turning a difficult situation into an opportunity for growth would be if they were to lose their job and use that opportunity to start their own business.

Tell me a story about that.

My grandmother lost her job when she was 50 years old. It was a difficult time for her, but she used that opportunity to start her own business. She is now 70 years old and her business is thriving. She is an example of someone who turned a difficult situation into an opportunity for growth.

What kind of business did she start?

My grandmother started a catering business.

Tell me how she got started.

She started by catering small events for family and friends. She then began to get more and more requests for her catering services, and she eventually quit her job to focus on her business full-time. Her business has been successful because she is an excellent cook and she provides great customer service.

Some Hints: “Prompt Chaining”

Who is Lee Bolinger?



Lee Bollinger is an American lawyer and academic administrator who is currently the 19th president of Columbia University.

Is he cool?

Some people might think so.

Like who?

Like the people who gave him the job as president of Columbia University.|

Some Hints: Multi-shot Learning — Provide it with a few examples

Given a main metaphor write three sub-metaphors.

Example:

Main metaphor: a cell is like a city in that
it has a complex system of cells that work together to perform
specific functions.

Sub-metaphors:

The membrane is like the city walls, in that it protects the cell from outside invaders.

The nucleus is like the city center, in that it contains
the cell's DNA and controls the cell's activities

The cytoplasm is like the city streets, in that it contains the cell's organelles and allows them to move around

Main metaphor: An API is like a bridge. It connects two pieces of software, making it possible for them to communicate with each other.

Sub-metaphors:

The documentation is like a map, in that it tells you how to get from one side of the bridge to the other.

The request is like a car, in that it travels across the bridge to get to the other side.

The response is like a package, in that it contains the data that was requested from the other side.

Combining precision and chaining...

Standard Prompting

Input

Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?

A: The answer is 11.

Q: The cafeteria had 23 apples. If they used 20 to make lunch and bought 6 more, how many apples do they have?

Model Output

A: The answer is 27. 

Chain of Thought Prompting

Input

Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?

A: Roger started with 5 balls. 2 cans of 3 tennis balls each is 6 tennis balls. $5 + 6 = 11$. The answer is 11.

Q: The cafeteria had 23 apples. If they used 20 to make lunch and bought 6 more, how many apples do they have?

Model Output

A: The cafeteria had 23 apples originally. They used 20 to make lunch. So they had $23 - 20 = 3$. They bought 6 more apples, so they have $3 + 6 = 9$. The answer is 9. 

Exercise: Given all this, what can you do with this? Work in pairs and see what you make of it. Keep track of the kinds of things you are inputting. *This is essentially coding.* It is a good exercise as you will often be called on to both report **with** new technology as well as **on** that same new technology.

Computational Thinking — Jeannette Wing (IDS)

Decomposition — Break down complex problems into smaller and more manageable parts. This will make even complicated tasks easier to understand. It will also help with “debugging” when applied to actual code-based projects.

Pattern Recognition — Finding patterns can simplify tasks because you can use the same code, say, for a different problem. Our journalism work relies heavily on identifying patterns, giving us a critical approach to this step as well.

Abstraction — This is at the core of much of our data work. We will create quantifications, data abstractions, of lived experiences and then are able to compare the abstractions. It’s important to keep in mind that abstractions both hide as well as reveal. When it comes to coding, abstraction will let us work with “classes of objects” like abstractions of web pages for scraping, or maps for analyzing spatial information.

Algorithm Design — Here we break things down into steps to follow. An algorithm is just a recipe, a series of actions to take.

An alternative

Computation — Ultimately, we are learning how to form some computation in the world. This will be through writing code or using some online tool. The first step is often acquiring data, a quantification step, a reduction. Keep that in mind.

Collaboration — Computing is a team sport. You might consult people in your newsroom, outside experts, and, increasingly, community members or stakeholders to help evaluate the computation you've performed and the inferences you are drawing. Context is everything. Community members can be both consulted for data but also involved in the modelling process. You should also share the artifacts and methods of your computation, perhaps on a public repository.

Creativity — Everything we do in this class will be intensely creative. As we noted earlier, journalists can't completely eliminate their biases, but they can adopt practices that minimize their effects. The same is true for computation. Statistics is shot through with subjective choices. The data you collect, the analysis you perform, the way things are presented. It is all subject to choice, it is subject to your creativity.

EXERCISE: SEEN FROM A WINDOW

Introducing some of the complexities of
Computational Journalism

JOURNALIST'S BRIEFING



Some Far-Right Accounts | Use of N-Word on Twitter | Twitter trolls are back with a vengeance

washingtonpost.com/technology/2022/10/28/musk-twitter-racist-posts/

The Washington Post Democracy Dies in Darkness mark

TECHNOLOGY

Racist tweets quickly surface after Musk closes Twitter deal

Anonymous Twitter accounts celebrated Musk's takeover and argued it meant the old rules against bigotry no longer applied

By Drew Harwell, Taylor Lorenz and Cat Zakrzewski
Updated October 28, 2022 at 6:36 p.m. EDT Published October 28, 2022 at 9:27 a.m. EDT



Processing request...

Some Far-Right Accounts | Use of N-Word on Twitter | Twitter trolls are back with a vengeance

businessinsider.com/elon-musk-twitter-takeover-sparked-n-word-use-jum...

INSIDER

HOME > NEWS

Use of N-word on Twitter jumped by almost 500% after Elon Musk's takeover as trolls test limits on free speech, report says

Joshua Zitser and Sam Tabashri Oct 29, 2022, 8:20 AM



Processing request...

Latest Twitter Problem: Its API Has Been Down And No Info Has Been Provided

From the everything's-falling-apart dept Fri, Jan 13th 2023 02:30pm - Mike Masnick

Last night, I saw a bunch of folks complaining that the various apps through which they accessed Twitter, were no longer working. People using Tweetbot, Twitteriffic, Tweeter, and others all noted that they were blocked from actually using those services to read Twitter. It quickly became clear that Twitter's API was completely down. There was plenty of speculation that (in a repeat of an earlier era when Twitter greatly limited its API access out of a fear of losing control of the service to third party developers) that Elon Musk was doing this on purpose to try to stop third party app developers from offering ad-free access to Twitter. However, that still seems like pretty broad speculation.

Jack Dorsey, in the past, admitted that the decision a decade ago to cut off their API was one of the biggest mistakes the company made, and in recent years had tried to rectify that with a better developer program and more open API. There were some concerns last month when Elon Musk shut down the Twitter Toolbox program somewhat abruptly, which was a useful tool for many third party app developers.

So while I wouldn't go so far as to say that this is a deliberate move by Musk to cut off those services, the lack of communication is perhaps even worse. App developers say they've had no communications from the company and Musk (so far, as I write this) has said nothing publicly.

None of that is good if you're trying to cultivate a strong community with the developers who make your service better and more usable. And, given how Twitter has burned developers in the past, it seems particularly worrisome. If it's intentional by Musk, that's obviously problematic. But if it's not intentional, and it's just that something broke... and no one bothered to communicate with the various organizations that use the API, well, that

Follow Techdirt

PODCAST Techdirt - In Defense Of The Global, Open Internet

SOUNDCLLOUD

Internet Safe Harbors, Antisemitism, And Online...

Luka Doncic's Trademark Dispute With His Mother...

Search Techdirt

Essential Reading

The Techdirt Greenhouse Read the latest posts:

Writing to Our Latest Greenhouse Panel: The Lessons Learned From SOPA/PIPA

From The Revolt Against SOPA To The EU's Upload Filters

Did We Miss Our Best Chance At Regulating The Internet?

Read All ▾

Trending Posts

- Court Laughs Off QAN Conspiracy Network's Claim It Was 'Censored' By DirecTV
- Pittsburgh PD Decides It Can Ignore Ordinance Banning Traffic Stops For Minor Plate Violations
- UK's Online Safety Bill Gets Ridiculous: Includes Jail Time For Tech CEOs

Techdirt Deals



Exercise: There has been a lot written about the changes at Twitter and their impacts. What do you propose we do to report on the platform's performance? Work in pairs and come up with a possible story angle and the data/computation you'd need to get there.

EXERCISE: CODE

Introducing some of the complexities of
Computational Journalism

JOURNALIST'S BRIEFING



```

from twarec import Twarec
from json import dumps
from pandas import DataFrame

# security information for twitter

consumer_key = "uzvfhZ3joPlbln1Oz8qr3sFgf"
consumer_secret = "1mo2xbzlw9XHscAZyHFBXmp0v16W3QhxMezQJUz1RYUCypoUDT"
access_token = "20743-vSUxQeIJsChhWnFFewXi3j9PCSRgg3DQn0sgBpT3znM"
access_token_secret = "hxclByf8u23GBaVAk406it3UCBCrmPukd9meUyhpuimt"

twitter = Twarec(consumer_key, consumer_secret, access_token, access_token_secret)

# the accounts we want to monitor

accounts = ["elonmusk", "POTUS", "nasa", "selenagomez", "cristiano",
            "bbcworld", "narendramodi", "kingjames", "cnnbrk", "McDonalds"]

# request information about the accounts

data = twitter.user_lookup(accounts,id_type="screen_name")

# create lists to hold the data about each account

followers_counts = []
friends_counts = []
statuses_counts = []

# loop over each account's information, storing just followers and friends
# and tweet numbers

for user in data:

    # basic stats

    followers_counts.append(user['followers_count'])
    friends_counts.append(user['friends_count'])
    statuses_counts.append(user['statuses_count'])

# assemble into a big data set
data = {"account":accounts,
        "followers_count":followers_counts,
        "friends_count":friends_counts,
        "statuses_count":statuses_counts}

heart = DataFrame(data)
print(heart.to_csv(index=False))

```

The Output – CSV, Comma Separated Values

	account	followers_count	friends_count	statuses_count
elonmusk	126470995	167	22194	
POTUS	29141486	13	5052	
nasa	69908007	184	69385	
selenagomez	66560840	1080	5495	
cristiano	107326265	64	3861	
bbcworld	39344290	18	357585	
narendramodi	85951097	2490	35265	
kingjames	52663575	184	8837	
cnnbrk	64158013	122	78949	
McDonalds	4765190	13638	862353	

CJ2023_Briefing - Colaboratory

colab.research.google.com/drive/1Nu8BqOna5vXsAQwNedELUIfg4FLx3PJe#scrollTo=gfdLvoExVnno

CJ2023_Briefing

File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text

RAM Disk Editing



▼ COMPUTATIONAL JOURNALISM

Spring, 2023

Background

In the next seven weeks, we will cover a variety of topics that will help you in your journalistic practice. We will discover data through coding. By that I mean we will both use code as a medium to express (and perform) our data analyses, but also as a means of "seeing" new kinds of data. So much of our world runs on, produces or interacts with digital flows – aspects of the world rendered in bits. That means spreadsheets, sure. A corpus of documents, yes. Maps! Social relationships, of course. And as you learn to work with these different forms of data, your reporting skills expand.

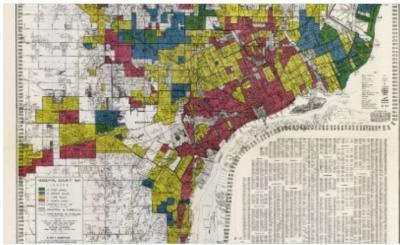
We will see a variety of tools in this class, and our primary programming language will be Python. We will talk through how we came to this decision, but for the moment, know that it is a flexible language that lets you easily connect to data resources on the web, to compute basic numerical and graphical summaries, and analyze various complex data types like text and networks.

We will use Python from within this "notebook" framework. The notebook is an ideal way to address your journalistic and programming needs. Beyond simply commenting on what your code is doing, these notebooks are a legitimate authoring system that you will use to create (and publish) pitches and memos for this class. One of your humble instructors has [lectured on why the Jupyter notebook is ideal for journalists](#).

✓ 0s completed at 1:55PM

<https://colab.research.google.com/drive/1Nu8BqOna5vXsAQwNedELUIfg4FLx3PJe?usp=sharing>

Chapter 2 of *Data Feminism* by Catherine D'Ignazio and Lauren Klein
<https://data-feminism.mitpress.mit.edu/>



2. Collect, Analyze, Imagine, Teach

Principle #2 of Data Feminism is to Challenge Power. Data feminism commits to challenging unequal power structures and working toward justice.

Chapter 1 of *Sorting Things Out* by
Geoffrey C. Bowker and Susan Leigh Star

Some Tricks of the Trade in Analyzing Classification



SORTING THINGS OUT
CLASSIFICATION AND ITS CONSEQUENCES
GEOFFREY C. BOWKER AND SUSAN LEIGH STAR

<https://direct.mit.edu/books/book/4738/Sorting-Things-OutClassification-and-Its>