



GPT-2 Journalism: Can AI produce Mike Royko's writing?

By Teddy Kamin with help from Jon Chun and Katherine Elkins
Kenyon College IPHS 300



Introduction

To begin, newspapers are vital for the Fourth Estate which means that newspapers are the check to the checks and balances of government. The past decade has not been a very good time for newspapers, as they have been caught in the digital revolution. More and more newspapers are shutting down because of online competition, drop in ad revenue, declining circulation among young readers, and the hunger for the fastest news, 24/7. Newspapers and their effect on society are indispensable. Newspapers, staffed by hard-working, truth-seeking reporters who aren't afraid of doing whatever it takes to bring justice, earned the title "Fourth Estate" for a good reason.

Mike Royko was one of the all-time great journalists in America during the 20th century. Royko was a vital part of people's daily lives. He was the best newspaper columnist this city had ever known. He started writing a column at the *Chicago Daily News* in 1964, and when that paper folded in 1978, he moved to the *Chicago Sun-Times*, and then to the *Chicago Tribune* in 1984, where he wrote until his death.[1] Coming from the northwest side of Chicago and a bar-owning Polish family, he was a Chicagoan through and through. Royko was famous for his fictitious character/alter ego, Slats Grobnik, a stereotypical beer-drinking, sports-loving, middle-class white ethnic Chicago male. Royko would never pull punches or tone down his words; his blunt edge sometimes ruffled readers' feathers. His blue-collar commentary was able to perfectly capture the Windy City and all of its characters. After his death, fellow Chicagoan and author Studs Terkel said, "If somebody says, 'What was Chicago like in the last half of the 20th century?' you'd say read Royko. He captured the city like no one else has ever captured a city and Chicago was his metaphor for the rest of the country." Most kids my age from Chicago have no idea who Mike Royko was. Growing up a newspaper kid, I have always heard about my parents' experiences with Royko around the *Tribune* newsroom and all of his lore. Royko was a life-long Cubs fan and his last column (he didn't know it would be his last) was coincidentally about the Cubs. I thought it would be fitting to use the magic of GPT-2 to see what Royko would have said about the 2016 World Series. In an interview with the *Chicago Tribune*, his widow, Judy, said, "The Cubs winning the World Series? There really are no words to describe what Mike would have thought, would have felt. He would have ... I don't have the words." Well,

Methodology

In order to generate an article in Mike Royko's writing style, a 355M-sized GPT-2 model was fine-tuned to a corpus of all of his columns from 1996-1997. He wrote pretty much every weekday, so this was a sample of around 450 articles of his writing from the *Chicago Tribune*.

I ran into some trouble cleaning the code for the first time of the text file and this generated some really odd text that wasn't coherent. Professor Chun was able to clean the text properly for me and then get 50 samples from that. These 50 samples were generated from a custom prompt.

During the training process the text of these articles was fed to the GPT-2 model 5,000 times, with the model learning the Royko style of witty commentary and short punchy sentences. Following this, the model used its understanding of the patterns contained in human language combined with the specific writing style of Mike Royko. Then the model was fed a specific prompt to write about. I thought it would be interesting to see what Royko would have written about the Cubs winning the World Series if he were alive. Royko being a life-long suffering Cubs fan, and a subject he often would write (er, complain) about in his column, I thought this would be a great fit. Also, the other part of the prompt I gave the model was about Barack Obama, a Chicagoan, being elected as the first black President of the United States. The temperature for all the text generations was set at 0.7 to keep the topic more focused. The higher the temperature the crazier the text becomes.

What is GPT-2?

GPT-2, created by OpenAI, is a transformer-based language model with 1.5 billion parameters, and it has been trained on 8 million webpages with the goal of accurately predicting the next word in a piece of text, given all the previous words in that text [2]. This goal makes GPT-2 capable of completing various tasks: It can answer questions about the content and comprehension of texts, summarize, and translate. It can also generate synthetic texts in which it mimics the style and content of a given input [2]. In order to mitigate possible malicious use and provide additional time to analyze the implications of releasing GPT-2 1.5BM, OpenAI is taking a staged-release approach, with the most powerful version currently available being GPT-3. In addition, GPT-2 outperforms other language models trained on specific domains (like Wikipedia, news, or books) without needing to use these domain-specific training datasets. On language tasks like question answering, reading comprehension, summarization, and translation, GPT-2 begins to learn these tasks from the raw text, using no task-specific training data.[2]

Metrics

Chicago Vernacular-Royko's signature style was to write in the voice of an everyday Chicagoan, specifically from the blue-collar Bungalow Belt. Always taking potshots at the powerful, and sticking up for the little guy.

Sentence Length-Short and punchy.

Scruffy Tone-He was not the friendliest character around town, the sort of guy who grumbled into his beer mug.

Wry Humor-You could always find yourself laughing out loud at the end of a Royko column.

Profound Wisdom-Royko had a knack for packing a serious punch out of the blue. You might be chuckling along at his wry humor and all of a sudden you found yourself struck by his power of his wisdom.

Reductionism-A signature move of Royko was his knack for constructing sentences so they ended with a rhetorical cymbal crash.

Results

Prompt: I never thought I'd live to see the day both the Cubs win the World Series and a Chicago politician like Barack Obama is elected President. Who could've imagined?

"You always write about the Cubs," Slats Grobnik said, "and I agree with you 100 percent." You really should, Joe, and this is just an example of why you shouldn't. "Just an example, Joe. Would you please put this on your column sometime in the next week or so." I thought so. "Right now, I'm thinking about putting you in a boa and jewelry case and burying you in the backyard." **From Sample 1**

"But the fans did. And it wasn't just the players. Mike - the old gutsy guy - became a saint. The other players followed suit. Eddie - the boobie - became a saint. And the fans have never been the same since. I wonder - is there anything that can bring a tear to the eye of Barry, the big, goofy, old-fashioned guy? And don't forget to ask Sid the elephant, Al, how he feels. He was the best pitcher in the world, a two-time defending champion. And you ever see the old timer, Slugger, fly out of the bullpen on a spring day in 1945, accompanied by the best band in the world? Willie Mays, the other great Cubs pitcher, is buried in St. Charles. So there you have it - a complete package deal. What's not to like? Especially when you get a 20-year anniversary next spring. I figured I'd save you the bother and just say that you can send your own personalized thank-you note. I'll even give you a free ride on my elevator." **From Sample 2**

"I never thought I'd live to see the day both the Cubs win the World Series and a Chicago politician like Barack Obama is elected President. Who could've imagined. But it seems like only yesterday that those Cubs fans were chortling and bemoaning the naiveté of a city kid named Art Buchholz, who thought he could hijack the Cubs franchise and challenge the white ownership structure. The Cubs were a cute team at the time, but they were mediocre. You didn't win championships by mid-summer dropping Fuhrman on your back and rolling over and making him say: "Ha, ha! Finally the rival Cubs are beating up on us, and we're winning." That was the last we heard of Art Buchholz. And it was obvious that he was a very smart guy in the city kid's position. He had the education and opportunity to go somewhere else if he had the stomach to. But he naively put all those skills to use running a pizza parlor. What a guy. What a rookie. There was talk in some of the other papers about what a guy Art Buchholz was and how dumb he was. I didn't know that he was just a little bit goofy. What about you, Art? Might I have omitted your name? By the way, it is possible that Art Buchholz is named after an old millstone fashioned around a crooked nose. If you want to be on the winning side of this argument, send Art a check or call 1-800-636-3363 and tell him what you think. You might just surprise him with how goofy people can be." **From Sample 10**

You could've said: "See, you have a Chicago wedding cakewalk, followed by a champagne toast in the Abbey bar, followed by a kissy-face with the groom on the cheek, and then a kissy-face with the bride at the nuptials. . . . Ah, honey, you have created a toast that is both charming and alcoholic." You could've done worse, I suppose, except for the cakewalk and kissy-face. But I wish I could've done better, because . . . no, that's unfair. I should've been praised for at least trying. Now I just don't get it. What is the appeal of a wedding? Some people say it gives them an opportunity to experience life as an individual. Is that bowling, golf, tennis or even basketball? The groom said: "No, a golf bowling club. . . . Imagine rolling in the parking lot after work and saying to yourself: 'Oh, joy, fulfillment, fulfillment.'" That's what I've always said about weddings. Maybe it's the music. Or the flowers. But I still feel that we have

Conclusion

Writers across America have tried to write like Royko with his scruffy tone and signature reductionism that makes his writing like no one else's. Many would-be Roykos have made it into print after attempting the feat, but none have come close to the real deal. He is an expert at social commentary, able to penetrate the veneer of the everyday, whether the political scene or nitty-gritty city life, with an uncanny knack for exposing the otherwise unspoken truth. He puts in writing what everyone else is afraid to say. Royko turns everything on his head. No one will ever reach the same level of writing genius as Royko, but GPT-2 gave him a good run for his money.

In the samples, the first thing that I was interested in seeing was whether the generated text would have been able to pull off the wry humor and scruffy tone that is iconic to Royko.

In the first sample, I was glad to see GPT-2 was able to identify Slats Grobnik as Royko's alter ego and give him snappy dialogue. Slats, who was once described by one of Royko's colleagues as "an urban Huck Finn who lives in a three-flat, hangs out in a tavern, cracks his knuckles," happened to sound just like the writer who birthed him into newspaper lore[3]. It is very impressive that AI-Slats was able to crack a Grobnik-esque joke about burying this guy Joe in a boa and jewelry case. Classic wiseguy.

In the second sample, I found a signature case of Chicago vernacular when referring to someone in the story by only their first name (a textbook Royko move) and making him just an everyday guy. GPT-2 was able to describe this character in a way that totally made me believe this guy could be a real Chicagoan.

In the third sample, AI-Royko took its turn to make fun of Art Buchholz who tried to hijack the Cubs. It created this wonderfully funny backstory about the character, and even prompted the reader to send Art a note or call him and let him know how you felt. His wry humor was in full effect here, giving this guy plenty of grief and putting him on blast in the paper, which is another thing that is *completely* fitting to the real Royko.

This final sample might have been my favorite and most similar to something you would actually see in the paper. A classic case of Royko's reductionism, stripping something so extravagant as a wedding down to the cake. He is able to poke fun at the cultural concept of weddings through messing with bride and groom and all the "kissy-face" this and that. His final zinger, at the end of the last sentence: he concedes there is one good thing about a wedding, and that's the cake. Once again, AI-Royko is able to be proficient in one of Royko's iconic writing techniques.

Overall, I think that GPT-2 was extremely effective at achieving a very similar writing style to Royko. It was able to combine his reductionism and wry humor while still delivering penetrating wisdom, all while talking about the Cubs having won the World Series. Another thing I was impressed by was GPT-2's ability to talk about specific sports players

Sources/Acknowledgements

[1] <https://www.chicagotribune.com/opinion/commentary/ct-mike-royko-20-years-death-perspec-0416-20170414-story.html>

[2] <https://openai.com/blog/better-language-models/>
[3] <https://www.chicagotribune.com/news/ct-xpm-2001-06-13-0106130008-story.html>

Big thanks to Professor Chun for doing all the coding and generating the text! Both Professor Chun and Elkins have been teaching us amazing concepts that are really cool and useful in the real world!