

The Rise of The Bots

How do automatically generated articles differ from human-written articles?

Filippa Heierli, University of Fribourg (CH)

Daniel Beck, University of Fribourg (CH)

2. Jahrestagung der Fachgruppe
Sportkommunikation und Mediensport der DGPK
Leipzig, 27. September 2019

ALGORITHM

BOT

COMPUTATIONAL JOURNALISM

AI in Journalism

The Associated Press announced in an advisory to customers today that the majority of U.S. corporate earnings stories for our business news report will eventually be produced using automation technology.

Source: The Associated Press (2014)

Fußball

Roboterjournalisten in der Kreisliga

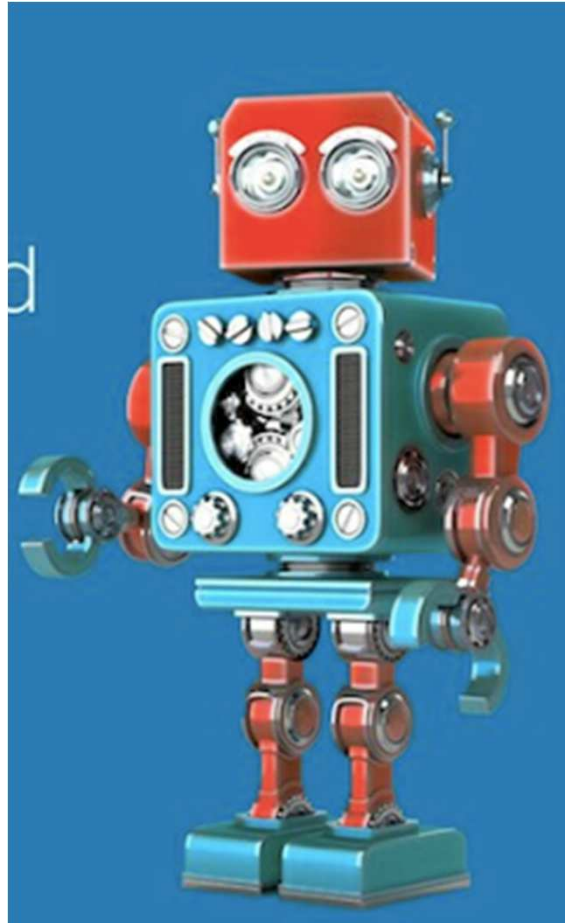
Der DFB initiiert zur neuen Saison im Amateurfußball das größte Roboterjournalismus-Projekt weltweit. Aus den reinen Spielberichtsdaten aller 70.000 Begegnungen pro Woche wird für jedes Spiel ein Artikel generiert – von einer intelligenten Software.

Source: Deutschlandfunk.de (2019)



Source: The Washington Post (2016)

AI in Journalism



Source: The Washington Post (2017)
Slide 6

Example - Article by Heliograf

Wilson at Yorktown
Football · August 31, 2017

Wilson 14-20 **Yorktown**
Overall: (9-6) Conf: (1-1) Overall: (9-2) Conf: (4-1)

	1	2	3	4	Total
W	0	0	7	7	14
Y	0	7	0	7	20

Game Summary:
The Yorktown Patriots triumphed over the visiting Wilson Tigers in a close game on Thursday, 20-14. The game began with a scoreless first quarter. In the second quarter, The Patriots' Paul Dalzell was the first to put points on the board with a two-yard touchdown reception off a pass from quarterback William Porter. Wilson was behind Yorktown 7-0 heading into the second half. Wilson's Anton Reed tied the score with a two-yard touchdown run. The Patriots took the lead from Wilson with a two-yard touchdown run by Tanner Wall. The Patriots scored again on Adam Luncher's 29-yard field goal. Yorktown maintained their lead going into the fourth quarter, 17-7. The Patriots extended their lead over the Tigers on Luncher's 27-yard field goal. Wilson cut into the Patriots' lead with a three-yard touchdown run by Amir Gerald. The game ended with Yorktown defeating Wilson, 20-14. Yorktown's top passer was Wilson, who completed 6 of 10 passes for 91 yards and one touchdown. Yorktown's top rusher was Wall, who had seven carries for 57 yards and one touchdown. The Patriots' top receiver was Wall, who had four catches for 54 yards. Yorktown will play Wakefield High School (1-0, 0-0) on Sept. 8. Wilson will play McKinley Technical High School (0-1, 0-0) on Sept. 8. This story may be updated if more information becomes available. It is powered by Heliograf, The Post's artificial intelligence system.

Game Summary:

The Yorktown Patriots triumphed over the visiting Wilson Tigers in a close game on Thursday, 20-14. The game began with a scoreless first quarter.

In the second quarter, The Patriots' Paul Dalzell was the first to put points on the board with a two-yard touchdown reception off a pass from quarterback William Porter.

Wilson was behind Yorktown 7-0 heading into the second half. Wilson's Anton Reed tied the score with a two-yard touchdown run. The Patriots took the lead from Wilson with a two-yard touchdown run by Tanner Wall. The Patriots scored again on Adam Luncher's 29-yard field goal.

Yorktown maintained their lead going into the fourth quarter, 17-7. The Patriots extended their lead over the Tigers on Luncher's 27-yard field goal. Wilson cut into the Patriots' lead with a three-yard touchdown run by Amir Gerald. The game ended with Yorktown defeating Wilson, 20-14.

Yorktown's top passer was Wilson, who completed 6 of 10 passes for 91 yards and one touchdown.

Yorktown's top rusher was Wall, who had seven carries for 57 yards and one touchdown. The Patriots' top receiver was Wall, who had four catches for 54 yards.

Yorktown will play Wakefield High School (1-0, 0-0) on Sept. 8. Wilson will play McKinley Technical High School (0-1, 0-0) on Sept. 8.

This story may be updated if more information becomes available. It is powered by Heliograf, The Post's artificial intelligence system.

Source: The Washington Post (2018)

State of Research

- Readers consider software-generated content to be **more informative, accurate, objective and trustworthy** than human-written articles (Clerwall 2014)
- Readers consider the **credibility** of articles written by bots and journalists **as equal** (Van der Kaa & Kramer 2014)
- Algorithmic generated content considered as **more boring** than human-written articles (Clerwall 2014)
- Journalists should **focus on their “human” strengths** instead of trying to directly compete with artificial intelligence systems (Van Dalen 2012: 653)

**In what way
do automatically
generated articles
differ from human-
written articles?**

Methodology

- Quantitative & qualitative content analysis
- Sample:
 - Time period: August 24, 2017– December 15, 2017
 - The Washington Post: 157 articles
 - LA Daily News: 39 articles

Quantitative Analysis – Dimensions of Quality

Formal Categories	
<ul style="list-style-type: none"> • Medium • Date of publication • Serial number • Title • Version of publication • Date of reported game 	<ul style="list-style-type: none"> • Matching dates • Name of author • Role of author • Number of characters • Type of article
Dimensions of quality categories	
<p><i>Variety</i></p> <ul style="list-style-type: none"> • Opinion • Number of actors • Type of actor • Number of actors quoted 	<p><i>Relevance</i></p> <ul style="list-style-type: none"> • Internal relevance • Up-to-date
<p><i>Impartiality</i></p> <ul style="list-style-type: none"> • News balance • Neutrality • Equal say 	<p><i>Appropriateness</i></p> <ul style="list-style-type: none"> • Source transparency • Author transparency • Quantities, times and figures

(Based on Voigt 2016)

Slide 11

Qualitative Analysis – Measuring Language Intensity

Category	Example
Noun	The players Trevor Gill, Kyle Bryant, and Nehemiah Thompson have been the ‘three-headed monster’ all season (LA Daily News; December 3, 2017; Article 32).
Verb	Tigers’ quarterback Andrew Frias polished off a one-yard touchdown run (LA Daily News; December 3, 2017; Article 32).
Adjective	He was a lightweight football player.
Exaggeration	It took the team forever to score a touchdown.
Repetition	They played a very, very bad game.
Others	An interception by Parker Seitz helped start the party and quarterback Connor Downs’ rushing touchdown was icing on the cake of Valencia’s 5234 home win that advanced it to championship game, its first since 2013 (LA Daily News; November 25, 2017; Article 21).

(Based on Hamilton & Stewart (1993), Mellado & van Dalen (2014))

Speed of Publication

- 39 articles vs. 157 articles (same publication time scale)
- Longer articles by human journalists
- < 40% of Heliograf's articles published on same day
- 100% of Heliograf's articles as «updated»

Flexibility and Creativity

- Title structure
 - Heliograf: «North Stafford at Osbourn»
 - LA Daily News: «Valencia football gets tricky in win over Chaparral»
- Sentence structure
 - Heliograf: «The St. John's Cadets defeated the DeMatha Stags in a close game on Saturday, 17–14»
- Surprising events / Uniqueness
- Human journalists use more creative elements

«... *quarterback Connor Downs' rushing touchdown was icing on the cake of Valencia's 52:34 home win...*»

Variety

- Human journalists share opinions
 - LA Daily News: > 70% contained opinions
 - Entertaining
- Heliograf adds to overall journalistic diversity
 - Produces more articles than human journalists

Relevance, Impartiality, Appropriateness

- R: Similar tendencies
- Both «journalist types» fulfill internal relevance: Who, what, when, where, why
- I: Heliograf exceeds in neutrality and news balance on micro level
- No opinions → balanced news / neutral → more boring
- A: Heliograf: complete author transparency
- A: Human-journalists show more transparency when it comes to actor names and functions

Conclusion

- Study compares articles written by human journalists with automatically generated articles
- Bots challenge human journalists in terms of **production speed** and **impartiality**
- Humans use **more and different types of creative elements**
- **Combination of bots and human journalists** in the news room
- Human journalists should focus on **setting topics into context** and **predicting**

Thank you for your attention!

f.heierli@gmail.com

daniel.beck@unifr.ch