

## **AWS Publishing:**

Publishing – the art of possible

1st Oct 2020

### Global publishers trust AWS with their most crucial workloads

Across content creation and acquisition, supply chain and content management, publishing/distribution and data analytics and monetisation



Newsweek



































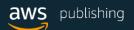
The Washington Post











## The reader is at the heart of the publishing transformation

Customer experience and engagement



Personalisation



Ubiquitous connectivity





Content consumption "on the go"





I think it's mission-critical that we continue to up our game as a digital product and technology company.

Meredith Kopit Levien, the upcoming CEO of the NYT





## "All-in approach" as a cornerstone of transformation and innovation

### ringier axel springer



### Challenges

- The publishing industry is changing rapidly, and Ringier Axel Springer (RAS) management wanted to be able to innovate and deliver products faster
- Management realized their private cloud can't compete with services offered by public cloud and wanted to refocus their developers to create digital products that drive the company revenue

### **Solution**

- RAS is migrating their 'private cloud' of applications constituting 80% of the company business to the AWS cloud. These applications are being re-platformed with some architecture redesign enabling RAS to realize cost benefit right after migration to the cloud
- The migration has started in September 2018 and will take 3 years to complete

### **Benefits**

- By using market standard solutions from AWS, Ring Publishing digital publishing platform is more reliable, accessible & attractive to clients
- The migration allowed RAS to free resources from infrastructure investments, tap into new fields of development and focus on critical topics like AI/ML
- Media brands of RAS were able to develop new features and stay flexible, even during the migration process and during COVID-19 crisis

Headquarters: Switzerland, presence across Central and Eastern Europe (CEE)

#### Website:

https://www.ringieraxelspringer.com/

#### **About RAS**

Ringier Axel Springer (RAS) is a leading digital media company in CEE, a joint venture between Swiss Ringier AG and German Axel Springer SE, created in order to combine their international operations in CEE.

RAS manages an extensive portfolio of more than 200 digital and print products in the expanding markets of Poland, Hungary, Slovakia, Serbia, Estonia, Lithuania and Latvia. Additionally, RAS has many on-line services VOD, classifieds, news, price comparison, e-commerce, and more.

Ring Publishing is an all-in-one digital publishing platform, helping media brands face challenges of the digital era.

H

Media today is not only a media game, it is a tech game and unless you are really very familiar with the big trends like AI, ML, you will not get the potential you need in order to make your business successful in the future. We went with AWS "all-in" and plan to migrate all our content services and big data stack to the AWS environment within 3 pears and already in the first 9 months we see significant progress.

Mark Dekan, CEO





https://www.youtube.com/watch?v=dbKdFB59YWE



# Disruption means changes to publishing business and operation models

### **Business Drivers**

**How AWS Helps** 

### Monetization

Acquisition and retention of subscribers/readers

New product launches

**2** Optimization

Cost optimization Editorial optimization

### **Customer Data Platform**

Digital Engagement tools
Personalisation
ML/AI

### AI/ML

Outsourcing mundane activities
Automation



## AWS offers solutions across all publishing workflows





## Deep dive today:

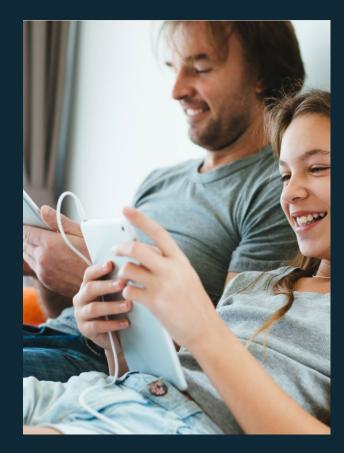
Understanding the reader: personalisation and recommendation



Innovation with AI/ML: Outsourcing mundane activities



Innovation with AI/ML: Optimisation





## Understanding every reader is needed in order to deliver unique experience

Today publishers collect data from ...

### **BILLIONS OF DEVICES**



Customers use multiple devices, browsers using anonymous / PII identifiers to interact with products, content and ads

## BILLIONS OF RELATIONSHIPS



Consumers view billions of web pages and apps, and generate billions of events across devices

## BILLIONS OF EVENTS THAT SIGNAL INTENT



Customer generate signals of intent such as search queries, product page views, ad clicks, purchases, and loyalty program enrollment



### What is a Reader Identity Graph?

### A Unified, 360-degree View Of The Customer



Persistent Ids across devices and channels

Audience building and segmentation

Cross-device frequency capping

Analyze customer journey



## Examples of the publishing use cases – digital engagement

Confirm purchase subscription, podcast

Respond to social media, customer reviews, support transcripts, and other conversations

Analyze user behavior across channels

Breaking news alerts
/ sports results

Personalized notifications based on profile preference of content

Articles recommendations

**New content alert** 

Login/password

Identify likely churners and trigger responses

Content recommendations

**Convert free users to subscribers** 



Single view of channel engagement



## Recirculation of trending content with Data Analytics

### Challenges

 Needed to develop a platform that analyzed real-time clickstream events such as readership statistics, impressions, and page views for more than 300 global websites and apps. In addition, the publisher wanted to monetise the text-to-speech feature

### **Solution**

- Hearst has partnered with AWS and Built a clickstream analytics platform that transmits and processes more than 30 terabytes of clickstream data a day, streamed from more than 300 Hearst websites worldwide
- · Amazon Kinesis Firehose automatically moves buffered data from Amazon Kinesis Data Streams into persistent storage on Amazon Simple Storage Service (Amazon S3)

### **Benefits**

- Fast insights With the clickstream analytics platform, Hearst can make the entire data stream - from website click to aggregated data - available to editors in minutes
- Increased content recirculation with the ability to get content metrics quickly, Hearst editors have increased the recirculation of trending content by more than 25 percent
- **Reduced complexity**
- Simplified data analysis

I don't know how we could have made our clickstream data pipeline work without Amazon Kinesis services



### HEARST

Headquarters: USA

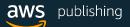
Website:

https://www.hearst.com/

### **About Hearst**

Hearst Corporation, headquartered in New York City, is one of the largest media and information companies in the world.

Some of the newspapers and magazines owned by them include Elle, Marie Claire, Cosmopolitan or San Francisco Chronicle.



## 1

## **Editorial personalisation: Targeted Content**

### Challenges

 The Globe and Mail was planning to launch a new application that enables its growing online readership to access stories and breaking news from mobile devices. And to increase reader engagement, it wanted to serve up targeted articles based on each reader's individual interests.

### **Solution**

- Personalized search engine was built by AWS's partner, Clear Scale
- The recommendation system uses several AWS services, i.e. AWS Kinesis, AWS Dynamo DB, Amazon EMR
- The customer also uses Amazon S3, Amazon EC2 Amazon SQS and Amazon CloudWatch

### **Benefits**

- AWS offered an efficient solution with The Globe being able to get the service to the market in just three months
- The solution was flexible and allowed the Globe to improve its mobile app over time
- AWS helped to dramatically increased reader engagement—with offering readers relevant content

As we work to improve the reader experience, AWS and ClearScale are helping us get there both quickly and cost-effectively.

-Mike O'Neill, Director of Data Science, The Globe and Mail

### THE GLOBE AND MAIL\*

Headquarters: Canada

Website:

https://www.theglobeandmail.com/

#### **About The Globe and Mail**

The Globe and Mail is Canada's most read newspaper with a national weekly digital readership of 4.7 million.

In print for 170 years, the newspaper delivers coverage of national, international, business, technology, arts, entertainment, and lifestyle news.



### Daily newsletter helps to generate traffic

Challenges	Solution	Benefits
<ul> <li>Daily Voice team was looking for a scalable infrastructure that could support its ambitions for growth and decided to go with AWS</li> <li>Additionally, the customer was also looking for ways of effectively engaging its digital readers and optimising the distribution of its daily newsletter</li> </ul>	In addition to AWS services used for optimizing and automating its environment (e.g. Amazon S3, Amazon EC2, Amazon CloudFront), the customer decided to deploy Amazon Simple Email Service (Amazon SES) on a daily basis	<ul> <li>Daily Voice typically sends around 200,000 emails a day through Amazon SES to its opt-in readers</li> <li>Daily Voice traffic patterns are now strongly tied to their daily email, a digest of the most popular stories that exist in that reader's town</li> </ul>



AWS was and is widely recognized as the market leader, and we recognized all these extra tools AWS provides that make administration so much easier. That's definitely been proven out in the time we've been working with AWS.

Within the last year, our traffic has grown about 40 percent, and we've easily been able to scale up.

It's not something we really think about, it's so simple.

— Travis Hardman, CTO



Headquarters: USA

Website: https://dailyvoice.com/

### **About Daily Voice**

Daily Voice, is an American community journalism company specializing in hyperlocal media, that is based in Norwalk, Connecticut, and currently operates a significant number of town-based news web sites in Westchester County, Dutchess County, Putnam County, Rockland County in New York; Bergen County, Passaic County in New Jersey; and Fairfield County, Connecticut. Currently, Daily Voice covers news in 43 different cities across three counties, brings readers the news they need about communities they love



### Examples of the publishing use cases – ML/AI

**Automated content creation** 

Automated social media distribution

**Audio** 

Articles recommendations

**Unlocking archives** 

**Content moderation** 



**Aiding reporting** 

Content monetisation strategies



## 2

## Outsourcing mundane activities:

## Auto-tagging Photos with ML

### **Challenges**

- Aller Group recognized that in order to increase reuse or sell their photo inventory, they had to significantly improve tagging of the photos
- The customer was looking for a quick and cost-efficient way to auto-tag 13 million photos that they have in their current inventory
- Manual work would require approximately 44 FTE to fulfil the task

### **Solution**

- Aller, who have several mission critical applications hosted on AWS, decided to use AWS Rekognition for images tagging (i.e. person detection, scene and activity keywords, etc.)
- AWS Comprehend was used for user comments (sentiment) and Reuters Open Calais was used for articles (keywords, concepts)

### **Benefits**

- Photo-tagging was successful and has led to a reduction in tagging costs from € 3.85 per image to € 0.12 per image
- The new solution runs 24/7 and allows for continuous improvements and multi-lingual tagging
- Aller was also able to easily experiment with different thresholds for facial comparison
- The customer has since decided to also extend the tagging to their video inventory

Using AWS Rekognition completely changed the way we see how Machine Learning can be implemented. The speed with which the PoC ran and the low costs associated with the overall solution has led us to reconsider several other projects that earlier had been seen as unrealistic.

- Thomas Culmsee Business Relationship Manager and Head of Editorial Production

© 2020. Amazon Web Services. Inc. or its Affiliates. All rights reserved.

### Aller

Headquarters: Denmark

Employees: 2.2K

Website: www.aller.dk

### **About Aller**

Aller Group is the leading magazine publisher in the Nordic countries, headquartered in Copenhagen, Denmark. It publishes among other Elle, Cafe, Familie Journalen, Femina, Allers and Se og Hør.

Today, Aller is the leading publisher of magazines and newspapers in the Nordic Region with a weekly circulation of 3.2 million copies.





# Improving auto-tagging with unsupervised, self-learning algorithm

### **Challenges**

- Aller Group has been using AWS Rekognition and AWS Comprehend for the photo tagging of their library since 2019 and has auto-tagged 5M photos in their current inventory
- The customer wanted to further innovate and decouple face detection and face clustering to eliminate false positives using other parts of Rekognition

### **Solution**

- Aller has partnered with Keycore, AWS's partner to build metadata generator system for Woodwing Elvis (can easily be expanded to other REST interfaces)
- The solution was built on top of AWS Rekognition as an unsupervised, selflearning, face clustering algorithm using a serverless (and massively parallel) version of DBScan
- It also works with video and identifies persons created via images

### **Benefits**

- The system has processed 5M images so far (Jul 2020) and has tagged 10M faces, 3M of which was discarded by the AI due to poor quality or position/size in the image
- The customer achieved improved tagging: With the solution, the number of images where the Crown Prince was tagged increased approximately 1000% picking up images previously missed where the Crown Prince is with other people.
- In addition, the solution resulted in cost and time savings due to automation of the entire process

"

Using AWS Rekognition completely changed the way we see how Machine Learning can be implemented. The speed with which the PoC ran and the low costs associated with the overall solution has led us to reconsider several other projects that earlier had been seen as unrealistic.

- Thomas Culmsee

Business Relationship Manager and Head of Editorial Production

### Aller

Headquarters: Denmark

Employees: 2.2K

Website: www.aller.dk

### **About Aller**

Aller Group is the leading magazine publisher in the Nordic countries, headquartered in Copenhagen, Denmark. It publishes among other Elle, Cafe, Familie Journalen, Femina, Allers and Se og Hør.

Today, Aller is the leading publisher of magazines and newspapers in the Nordic Region with a weekly circulation of 3.2 million copies.



## 3

## Mobility (smartphones & 5G) transformation: Al-powered News App

### **Challenges**

 Editorji required a cloud architecture that is faulttolerant and can scale rapidly to accommodate traffic spikes on the app and which enable smooth beta launch of the app in Oct 2018

### Solution

- To accommodate spikes in traffic, Editorji uses Elastic Load Balancing
- The customer also relies on Amazon Polly to voice its graphics-based videos and Amazon Comprehend for entity extraction and tagging in the cloud
- Additionally, it also uses Amazon Elastic Transcoder to convert media files from their source format into versions that will play back on mobile devices

#### **Benefits**

- Scales to accommodate 300% rise in application traffic
- Saves manpower, as infrastructure is managed by 1 person
- Automates tasks often performed by people, through AI and ML
- Enables expansion to new geographies



Headquarters: India
Website:
http://www.editorii.com

### **About Editorji**

Editorji is an Al-powered platform for the curation and distribution of video news. Founded in Delhi by Vikram Chandra, Editorji curates news clips to users' taste, making it possible to stream clips on demand. Audio and text formats are also available.



We strongly feel that when it comes to cloud computing, AWS is miles ahead of competitors.

— Biswajit Borkataky

Head of Operations and Platform Relationships, Editorji





## 3 Predicting revenue: ML leveraged to Determine which stories go behind paywall

### **Challenges**

### News Corp Australia for digital content, they have to determine which articles reside inside of a paywall, and which are publicly available

 The customer wanted to deploy a solution to help make its decisions around which article should go behind the paywall

### **Solution**

 The customer decided to use TensorFlow on Amazon SageMaker to predict the number of incremental subscriptions a news story will generate

### **Benefits**

- The solution helps the customer determine which news stories live behind a pay wall
- NewCorp is working on deploying the model in production and expects the model will enable editorial content producers to optimize the headline copy, the section where a story is placed, and other factors to maximize net new subscriptions generated for the business





Headquarters: Australia

Website:

https://www.newscorpaustralia.com/

### **About News Corp Australia**

News Corp Australia is Australia's number one media company, distributing news and information across a suite of digital and traditional print products.

News Corp Australia owns approximately 142 daily, Sunday, weekly, bi-weekly and tri-weekly newspapers, of which 102 are suburban publications (including 16 in which News Corp Australia has a 50% interest).



# **3 Speed&efficiency:** Using ML for matching images with news

**Challenges** 

 NewsUK needed to increase the search granularity and speed of finding a suitable image for a News story by using a more efficient content categorization and content search

### Solution

- The customer used Amazon
  Rekognition's pre-trained machine
  learning models via API's (labels,
  text, faces, celebrities, content
  moderation) to automatically
  analyze the existing images in their
  content hub
- This created new, rich metadata that contains detailed attributes about the discovered visual content.

### **Benefits**

- The customer is automatically able to categorize newly arriving image files, saving on cost and time
- News UK was able to create rich metadata for images in their content hub. This increases the speed and accuracy of identifying imagery for a given news story.





Headquarters: United Kingdom

Website: https://www.news.co.uk/

### **About News UK**

News UK is a British newspaper publisher, and a wholly owned subsidiary of the American mass media conglomerate News Corp.

It is the current publisher of The Times, The Sunday Times and The Sun newspapers; its former publications include the Today, News of the World and The London Paper newspapers.



## 3

## Aiding journalists: ML is helping journalists with reporting

### **Challenges**

- The Washington Post wanted to improve the work of their journalists by providing them with tools to break news faster
- In order to break news, their reporters need not just the words but also the structured data that is found in newsworthy documents

### **Solution**

The customer decided to use
Amazon Textract in order to equip
their journalists with machine
learning data extraction capabilities,
enabling them to study records of
public interest faster and to break
news based on insights from the
words and structured data of
newsworthy documents

### **Benefits**

- Amazon Textract reveals more data from documents, more accurately, aiding their journalists in their reporting
- Machine learning for document understanding is now a critical tool for The Washington Post to study records of public interest



### The Washington Post

Headquarters: USA

Website:

https://www.washingtonpost.com/

### **About Washington Post**

The Washington Post is an American daily newspaper published in Washington, D.C. It is the most widely circulated newspaper within the Washington metropolitan area. Daily broadsheet editions are printed for the District of Columbia, Maryland, and Virginia. The newspaper has won 69 Pulitzer Prizes.



## Globe & Mail use Amazon Polly for more authentic speech

#### **Solution Benefits Challenges** The Globe and Mail has explored · Amazon Polly mimic the workings of · Launched Audio Now where readers ways new technology could expand the human brain using Machine can listen to most of the content in a access and engagement of their Learning to build Neural-to-Text-torealistic newscaster voice journalism to audiences Speech systems (NTTS) • Built using Amazon Polly Newscaster • The systems generate sounds, then Available in English, French, and improves by listening to recorded Mandarin in male and female voices natural speech and imitating it · Allows the Globe and Mail to meet their readers, and listeners, wherever they are

Amazon Polly Newscaster enables us to provide our readers with more features to further their experience with our newspaper. This text-to-voice feature from AWS is miles ahead of anything we've heard to date.

— Greg Doufas

Chief Technical and Digital Officer



Headquarters: Canada

Website: www.theglobeandmail.com

### **About Global and Mail**

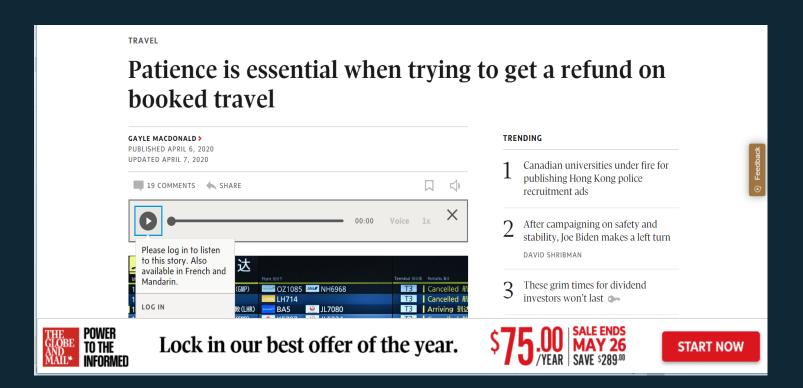
The Globe and Mail is a Canadian newspaper with print and digital versions reaching over 6M readers every week.

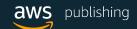
The Globe and Mail has won more national newspaper awards than any other news organization in Canada.





## Recent example – The Globe and Mail





# 3 Creation of audio books with text to speech

### **Challenges**

- Pearson is rapidly transforming their business from a print book business to a digital courseware, content and assessments business
- Pearson needed a technology partner to deliver a greater pace of innovation

### **Solution**

- Migration of strategic applications to AWS
- Amazon Polly for text-tospeech translation of a broad portfolio of books

### **Benefits**

- Creation of audio versions of textbooks at 1/10 the cost of paid voice actors
- Ability to immediately update audio textbooks each time editions are changed or an edit





Industry: Publishing

Headquarters: London

Employees: 32K

Website: www.pearson.com

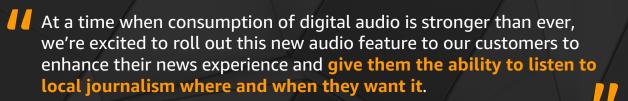
### **About Pearson**

Pearson provides education content and assessment, powered by services and technology, to enable effective teaching and personalized learning across 70 countries.



# 3 Monetization of text-to-speech feature

Challenges	Solution	Benefits
<ul> <li>McClatchy needed a solution to create audio versions of its articles</li> <li>In addition, the publisher wanted to monetise the text-to-speech feature</li> </ul>	<ul> <li>McClatchy partnered with AWS and AWS's partner, Trinity Audio, to generate audio versions of news articles within a matter of seconds</li> <li>The Trinity Audio technology also inserts advertising, resulting in a new revenue-generating stream</li> </ul>	Prior to the newsrooms-wide launch, the Al-powered feature was tested by the Sacramento Bee and the News & Observer in Raleigh for three months; each observed an increase in user engagement online. Time spent on each site increased by 168%, story page views went up 89% and visits per user increased by 95%.



- Jessica Gilbert, Senior Director of Product and Experience



Headquarters: USA

Website:

https://www.mcclatchy.com

### **About McClatchy**

McClatchy is and American publisher that operates 29 daily newspapers in fourteen states and has an average weekday circulation of 1.6 million and Sunday circulation of 2.4 million.

In addition to its daily newspapers, McClatchy also operates several websites and community papers, as well as a news agency, McClatchyDC, focused on political news from Washington, D.C



## 3 Polly leveraged for a book launch

campaign

**Challenges** 

- To support the release of Dan Brown's new novel, Origin, Doubleday, a division of Penguin Random House, asked BeeLiked to develop a digital campaign that would appeal to Dan's large social media fan base
- The use of the voice was important because the customer wanted to emulate the novel, which includes a personal assistant with a British accent.

### Solution

- Using Amazon Polly and its video platform, the customer created the original experience. After voting on a book cover, fans are transported to their own private book signing where Brian, Amazon Polly's British English male voice, welcomes them by name into Dan's workspace with the following words:
  - 'Welcome {first name}, we've been looking forward to your visit"

### **Benefits**

- · The video was shared widely under the hashtag #DanBrownOrigin and helped create a real buzz for the book launch
- The customer was impressed with Amazon Polly as a marketing tool, got a lot of positive feedback from users, and used it for another campaign

We chose Amazon Polly due to the ease of integration, the speed at which the voice file can be created, and the quality of the voices available.

> — Damian Dutton CEO and Founder, BeeLiked



**Headquarters:** United Kingdom

Website:

https://www.beeliked.com/

### **About BeeLiked**

BeeLiked is a digital marketing platform offering a wide range of campaigns to help brands engage with their existing audiences and reach new customers through the viral and social nature of the contests and games.



## Digital Innovation: Al-powered graph database helps users visualize connections

### **Challenges**

 Journalists and data scientists at The Globe and Mail needed to be able to quickly understand connections among millions of records to aid in investigative journalism and bring additional value to readers

### **Solution**

With the AWS Digital Innovation program the team learned how to work backwards from their customer, and developed an Al-powered graph database that connects the dots across people, events, organizations, etc. to help journalists see relationships that otherwise may not be apparent. It also enables readers to examine the data behind the stories – adding a level of transparency that had never been possible.

### **Benefits**

- Reduced project development time by 4 years
- Increased staff productivity
- Improved customer experience

The solution changes the [social] contract we have with our readers, and has a huge potential to change journalism. The biggest takeaway from the entire program is the idea that anything is possible. Where I would have spent five years persuading the team internally that we needed to create an entirely new data structure... it was done in under a year.

Editor-in-Chief David Walmsley

### THE GLOBE AND MAIL\*

Headquarters: Canada

Website:

www.theglobeandmail.com/

### **About The Globe and Mail**

The Globe and Mail is Canada's foremost news media company, reaching more than 6 million readers across print and digital every week.



## Interested in Amazon's approach to innovation? AWS Digital Innovation Program

Let's partner to invent something your customers will love.



1 EXPLORE innovation at Amazon – the principles behind our organization, architecture, mechanisms and culture



2 **DEFINE** a new product, service, or experience that will delight your customers; start with the customer and work backwards



BUILD a prototype in the cloud to bring your new digital offering to life; test it with your customers



## **Speedy subscriptions with Amazon Pay**

Challenges	Solution	Benefits
As journalism goes digital, The Washington Post has worked to become a more tech-focused company	The Washington Post integrated Amazon Pay into its website, allowing subscribers to pay with the credentials already stored in their Amazon account.	Faster checkouts - customers who purchase their subscriptions with Amazon Pay can check out in under 30 seconds
Washington Post was looking for ways to fuel its growth, including a speedier way to pay for its services		<ul> <li>High conversions – WaPo has boosted overall conversions on its site</li> <li>Fewer decline rates - Customers who sign up with Amazon Pay are declined just 2% of the time</li> </ul>

Our goal is to make the experience of signing up for a subscription as frictionless as possible – and that's what Amazon Pay offers

Miki King, Vice President of Marketing

## The Washington Post

Headquarters: USA

Website:

https://www.washingtonpost.com

### **About Washington Post**

The Washington Post is an American daily newspaper published in Washington, D.C. It is the most widely circulated newspaper within the Washington metropolitan area. Daily broadsheet editions are printed for the District of Columbia, Maryland, and Virginia. The newspaper has won 69 Pulitzer Prizes.

