

Premier Consulting Partner

PBS Offers Viewers Truly Personalized Experience with Amazon Personalize



Executive Summary

Public Broadcasting Service (PBS) is an Arlington, Virginia-based nonprofit organization founded in 1969 that broadcasts educational, news, and entertainment programs to more than 100 million television viewers across the U.S. and more than 32 million people online. PBS currently has approximately 330-member television stations, distributing the highest quality of content to all 50 U.S. states, Puerto Rico, U.S. Virgin Islands, Guam, and American Samoa.

PBS wanted to build a smart recommendation engine (SRE) capable of making high-quality suggestions to viewers based on a multitude of factors. To ensure success, PBS decided to partner with a cloud consultancy with AI/ML expertise and deep knowledge of the Amazon Web Services (AWS) platform. ClearScale was an excellent match for PBS and gave the nonprofit exactly what it needed to enhance viewer experiences significantly in the streaming age.



The Challenge

Like many of today's leading media and streaming platforms, PBS wanted to take its overall user experience to the next level. The organization hoped to provide audiences with better in-app programming recommendations based on a multitude of factors - deep links between titles, current popularity trends, user behavioral patterns, and more - to improve engagement and long-term loyalty.

On the surface, creating such a recommendation engine seems complex. Yet, the reality is that building these engines doesn't require data science expertise or AI/ML mastery. Companies only need to find the right combination of cloud-native tools and services, then feed them with their data. And with the right toolkit, these services don't take years to develop.

Fortunately, AWS offers managed AI/ML solutions that enable engineers to leverage pre-built models and automate much of the hard work of creating, training, and fine-tuning them. The challenge lies in knowing how to maximize what the cloud offers, especially given how quickly things change.

That's why PBS decided to approach ClearScale, an AWS Premier Consulting Partner with 11 AWS competencies, including Machine Learning, Nonprofit, and Data & Analytics. ClearScale is also a leader in MLOps, which is the type of technical expertise PBS needed to build the ideal recommendation system and sustain it over time. Together, PBS and ClearScale decided to move forward with an AWS-powered solution on top of Amazon Personalize.

"We worked with ClearScale to help set up and configure our initial solutions and data pipelines. We needed to leverage insights faster and launch something in months rather than years. Their experts set up an AWS cloud configuration and related services for using Amazon Personalize to help save us a tremendous amount of effort and thousands of engineering hours."

Director of Product Management, PBS



The ClearScale Solution

For PBS to build a truly differentiated recommendation system, the company needed the latest and greatest cloud technologies available, on top of expert implementation guidance.



Main Architecture Diagram

ClearScale came up with a detailed roadmap for tackling PBS's recommendation system project:

- Data Operations
- Machine Learning Operations
- Demonstrational User Interface



Data Operations

First, ClearScale and PBS determined together which data sources would feed into future ML models:

- PBS' Media Manager
- PBS' User Profiles
- Google Analytics metadata



Data Operations

Media Manager is a content management system that PBS member stations use to publish and share titles across different platforms. Media Manager also contains rich metadata, such as a product's release date, tags, author, etc., and comes with rules that contribute to deciding what gets shown to viewers in search results. For instance, Media Manager takes a viewer's age or location into account before making a recommendation. That way, young children don't accidentally come titles for older audiences or viewers in a region aren't recommended a news series from another location on the other side of the country.

PBS User Profiles contain valuable details about individual viewers, such as their previous interactions with PBS apps, their watchlists, watch times, and viewing history. Therefore, User Profiles contain some of the most obvious evidence of what people enjoy watching.



ClearScale and PBS also decided to incorporate contextual information from Google Analytics to gain a more comprehensive understanding of who watches PBS content and where. Google Analytics has non-sensitive data about people that can be useful in making inferences about their viewing preferences. The platform can also see what types of devices people use to watch content, which can serve as another data point for a recommendation system to consider. For example, a viewer might watch PBS news on her phone during her train commute to work. But, once at home, she might watch Daniel Tiger on TV with her kids.

To consolidate data from the first two sources, ClearScale set up a prototype environment for an Amazon Aurora for PostgreSQL relational database. The database existed in full isolation from PBS production systems to ensure maximum resiliency for ETL processes. Google Analytics data was captured via an ingestion pipeline and stored in Amazon S3.

ClearScale then implemented a data pipeline starting with AWS Glue, a serverless cloud-native solution to crawl, validate, and transform data from diverse sources. ClearScale also configured AWS Glue to make data consumable by formatting it into Parquet and offloading it into a data lake. These steps are all orchestrated using AWS Step Functions, allowing PBS to benefit from automated stateflow management and exceptions handling.

AWS Lake Formation and AWS Glue Data Catalog were crucial for securing PBS' data lake and pointing other cloud services to the right data stores. Data in the lake can be accessed in two ways, both using standard SQL:

- 1. Serverless analytics with Amazon Athena is best for ad-hoc exploration tasks when cost is the most crucial factor.
- 2. A robust data warehouse on top of Amazon Redshift for regular, well-defined queries with strict SLA requirements.

With the infrastructure for data operations in place, ClearScale was ready to tackle the MLOps side of the project.



Machine Learning Operations

ClearScale helped PBS establish the four primary stages of the ML lifecycle:

- Model development
- Training
- Inference
- Evaluation



Machine Learning Operations

Fortunately, AWS gives companies the ability to harness the power of data science and Machine Learning across these four stages without having to build models from the ground up.

ClearScale data engineers took on the task of creating the initial version of the smart recommendation engine based on Amazon Personalize while keeping in mind that PBS engineers would eventually take full ownership. ClearScale used Amazon FSx for Lustre to make data available for the system as it's loaded. The team also integrated Amazon SageMaker Studio as the development environment ML engineers would use to maintain models.



At the center of the model pre-production work are AWS Lambda, Amazon Athena, and AWS Step Functions. ClearScale connected them with Amazon Personalize to fetch data, load changes, and train the model. With these services in place, ClearScale selected the core recipes (which are Amazon Personalize algorithms fine-tuned for specific use cases) for PBS' SRE and built four models based on different requirements per recommendations input and output:

- **Popularity Count ML model:** suggests TV shows based on mainstream popularity. This is the simplest model in scope, yet it's important. Because other models dive deep into past data, they suggest programs relevant to the user yet distributed throughout history. In the Media & Entertainment industry, where the goal is to promote recent titles, this model helps others not go too deep into the weeds. By limiting the range of data taken into account to the previous week, it's possible to identify recent trends and augment them with predictions from other models. To keep those trends fresh, this model is retrained daily.
- Items Relationships ML model: suggests TV shows based on collaborative filtering to recommend programs that are most similar to ones the viewer interacted with before. This recipe (SIMS) digs deeper to reveal relationships between shows, including ones that are not evident to human intelligence at first glance, nor traditional linear and statistical algorithms.
- Interactions History ML model: suggests TV shows based on user behavioral patterns using active learning. With active learning we supply the model with user activities in the same session where we provide recommendations. This allows it to discover new rules in seconds without going through complete retraining, which would take hours.
- **Personalized Ranking ML model:** ranks TV shows based on apparent user preferences. Instead of fetching particular items, this algorithm takes ones supplied by PBS (e.g., "Best Christmas Shows" digest) and returns them in an order reflecting user preferences.

Machine Learning Models Comparison				
Criteria	Popularity Count	Items Relationships	Interactions History	Personalized Ranking
Patterns	Popularity	Similarity	Behavior	Behavior
Dimensionality	1,000s	10,000s	100,000s	10,000s
Performance	Best	Better	Good	Better
Coverage	Low	Medium	High	Medium
Accuracy	Good	Better	Best	Better
Retraining	Weekly	Weekly	Online Monthly	Weekly
Recipe	Popularity- Count	SIMS	User- Personalization	Personalized- Ranking



ClearScale deployed each of these models at Amazon Personalize's unified REST API, backed by the Amazon API Gateway, to make findings from PBS' recommendation engine available to the many platforms that support the company's streaming application. Access controls are based on Amazon Cognito and AWS IAM to ensure viewers only have access to their own data. Each model's API consists of four close-connected microservices:

- **Real-Time Recommendations API:** receives user information and, in a few seconds, offers recommendations on which great show will attract and entertain them next.
- **Personalized Notifications API:** do the same as the last microservice yet are used in conjunction with off-session marketing channels like SMS, email, or push notifications.
- **Feedback Loop API:** processes feedback from viewers in the form of "thumbs up" or "thumbs down" to determine their satisfaction with recommendations and, hence, their correctness.
- **Configuration Management API:** allows PBS administrators to fine-tune the recommendation engine on the fly without redeploying any system parts.

The world is not static in any sense, and neither is Machine Learning. As the environment evolves, trained models no longer operate as well as they did after being deployed. In 99% of cases, models degrade over time, decreasing the business value as well as end-user satisfaction: e.g., the items catalog receives new titles that are never seen by the model. In the best scenario, the model would refuse to recommend the title, introducing bias. In the worst scenario, the model would provide incorrect predictions leading to poor decisions. To ensure that the model is not frozen in place, it must be continuously retrained on the most up-to-date data and occasionally change its shape to fit new game rules.

The custom Model Monitor was added on top of Amazon CloudWatch to measure a precision metric that characterized the system's ability to make good recommendations to viewers. It doesn't just monitor metrics, it also makes automated decisions based on them (e.g., retrain the model when it's close to a certain threshold, so the metric value never drops below it, keeping viewers happy).

ClearScale's proof of concept for PBS yielded a "Precision at 10" metric of 0.0706. This number means, with every 10 titles recommended, at least one will be favored by the user with 71% probability. It's worth saying that many other recommender systems can only achieve a 0.03 result.



Demonstrational User Interface

The last phase of the project was to create a user interface prototype that would allow PBS viewers to personalize their accounts in a simple, visually engaging way. ClearScale created a demo web application that reused existing business logic and capitalized on the new recommendation engine.



Demonstrational User Interface

The Demo App was powered by **TypeScript**, **ReactJS**, and **Sass** for the UI, as well as for data management using **Effector** (client-side) and **React-Query** (API integration). While serving its purpose as a functional prototype, reflecting PBS's uniqueness with both styling and branding guidelines applied. And due to responsiveness, natively inherited from **Material-UI**, the demo app works equally well on desktops, tablets, and phones.

The demo user interface included the following components:

- "Web Hosting" to deliver the demo app to viewers and make it accessible regardless of platform.
- "Unified Auth" to allow PBS viewers to log in with existing credentials and automatically make their watch histories, preferences, and other personalization data available to SRE.
- A "**Title Card**" feature that displays details about a show when a person hovers over it in the catalog, as well as a rating indicating whether the title is relevant to the user.
- A "Content Player" to enable viewers to view recommendations in the demo app.
- A "**Top Picks for {User}**" feature that displays a personalized list to viewers based on the realtime recommendations API and its Interactions History ML model.
- A "Feedback Loop" feature that allows viewers to judge the relevance of recommendations



provided by the system and see in real-time how it affects the offered content.

• A "**Top {K} Over Last Week**" list that displays recent, popular titles across PBS' entire audience based on the Popularity Count ML model.



Top Picks for {User}



Title Card & Feedback Loop



Content Player



Top {K} Over Last Week



The Benefits

Today, PBS has an effective MLOps platform and recommendation system that it can build on going forward. The data pipeline that ClearScale set up cleans, validates, and enriches raw data PBS has accumulated over its 50-year history. The data that flows into the organization's recommendation system is consistent, accurate, and complete, making it a single source of truth for current and future AI-driven endeavors.

The new recommendation engine also gives PBS the ability to deliver more personalized experiences to viewers based on a myriad of factors. The four models that ClearScale built incorporate variables such as mainstream popularity, inter-title relationships, user behavior, and more to arrive at recommendations that are highly likely to please viewers.

Finally, the demo web application ClearScale developed for PBS showcases the power of the new recommendation engine in a user-friendly interface. It gives people the opportunity to quickly find titles they enjoy and share feedback on specific recommendations, empowering PBS to fine-tune viewers' experiences.

© 2022 ClearScale. All rights reserved. The ClearScale trademarks are the exclusive property of ClearScale, are registered with the U.S. Patent and Trademark Office, and may be registered or pending registration in other countries. All other trademarks, service marks, images, products and brands remain the sole property of their respective holders and do not imply endorsement or sponsorship.