

Towards AI-Driven Personalized Experiences in TV 3.0: An Extensible Personalization Architecture

Letícia Concolato Rangel

Marcelo F. Moreno

CITE THIS ARTICLE

Rangel, Letícia Concolato; Moreno, Marcelo F. ; 2026. Towards AI-Driven Personalized Experiences in TV 3.0: An Extensible Personalization Architecture. SET INTERNATIONAL JOURNAL OF BROADCAST ENGINEERING. ISSN Print: 2446-9246 ISSN Online: 2446-9432, doi: 10.18580/setijbe.2026.2 web link: <https://dx.doi.org/10.18580/setijbe.2026.2>



COPYRIGHT This work is made available under a Creative Commons Attribution-NonCommercial 4.0 International License.
Reproduction in whole or in part is permitted provided the source is acknowledged.

Towards AI-Driven Personalized Experiences in TV 3.0: An Extensible Personalization Architecture

Letícia Concolato Rangel and Marcelo F. Moreno, *Member, IEEE and SET*,

Abstract—This paper presents an extensible architecture for AI-driven personalization in TV 3.0 environments and its preliminary results. The proposed approach combines semantic analysis of audiovisual content with viewer-context information to dynamically select, instantiate, and adapt use-case templates for interactive applications. The architecture supports multiple personalization scenarios across live broadcasting and Video on Demand (VOD) services, enabling semantic and context-aware interactive experiences. Preliminary results demonstrate the feasibility of integrating content understanding, viewer-context modeling, and dynamic application composition within a unified personalization architecture, highlighting its potential to support more relevant and personalized TV 3.0 experiences.

Index Terms—TV 3.0; Artificial Intelligence; Personalization; Interactive Television; Semantic Content Analysis; Context-Aware Systems; User Experience

I. INTRODUCTION

THE growing adoption of digital platforms, such as streaming services and social media, has solidified a media consumption paradigm centered on personalization. Unlike traditional broadcast television, historically based on a one-to-many transmission model, these platforms employ recommendation algorithms and data-driven architectures. This infrastructure continuously analyzes user behavior, consumption history, and interaction patterns to deliver experiences tailored to individual preferences, transforming audience expectations regarding the discovery of and engagement with audiovisual content.

Attuned to this market reconfiguration, the Brazilian Terrestrial Digital Television System (SBTVD) Forum spearheaded the development of TV 3.0 [2]. This new generation of broadcasting transcends mere improvements in audiovisual technical quality, enabling interactive, dynamic, and highly personalized services. Through the native convergence of broadcast and broadband networks, TV 3.0 introduces technological capabilities for identifying active viewers and delivering contextualized content, services, and targeted advertisements, thereby enhancing communication relevance and audience engagement [1].

This pursuit of convergence addresses a long-standing debate regarding the future of media. A pivotal moment in this discourse occurred in 2015, when Reed Hastings, then co-CEO of Netflix, told *Wired* magazine that the linear television model, consolidated over five decades, had reached the end of its lifecycle. At the time, Hastings predicted the convergence of

television onto internet-connected screens, foreseeing that traditional channels would gradually be replaced by on-demand [6].

However, the evolution of the contemporary audiovisual ecosystem has unfolded into a scenario that is more complex and convergent than the definitive decline predicted at the time. Currently, the TV 3.0 specification reinterprets this paradigm by transforming free-to-air broadcast television into an experience that is natively app-based. Rather than succumbing to simple replacement by standalone streaming, TV 3.0 integrates mass broadcasting with broadband connectivity, converting channels into rich, interactive software ecosystems [7]. In this new landscape, the application evolves from a mere asynchronous catalog into a layer of real-time governance and personalization, enabling adaptive interfaces, contextual services, and dynamic engagement during linear programming playback [8].

Concurrently, Artificial Intelligence (AI) has become established within streaming platforms for automating curation, profile modeling, and semantic asset analysis. However, on such digital platforms, personalization mechanisms remain controlled by distribution intermediaries. In contrast, although broadcasters control the entire production and distribution chain, the design of interactive applications and contextual curation in television environments still relies predominantly on manual processes, creating a bottleneck for scalability and adaptability. Furthermore, unlike the deregulated web environment, broadcasters operate under strict editorial obligations, advertising regulations, and content rating policies. Therefore, introducing AI into TV 3.0 requires adaptive mechanisms that ensure regulatory compliance and broadcaster editorial control, positioning AI as an enabling technology for broadcast workflows.

Against this backdrop, this article presents an extensible architecture for semantic and context-aware personalization designed for TV 3.0 environments. The proposed approach employs AI techniques for the multimodal analysis of audiovisual content and the correlation between semantic descriptors and viewer context variables, enabling the dynamic instantiation and parameterization of templates for multiple use cases. Operating resiliently and agnostically across both broadcast and broadband networks, the architecture establishes the infrastructure needed to deliver AI-driven, interactive, scalable, and personalized experiences for the next generation of digital television.

Authors are with the Computer Science Department, Federal University of Juiz de Fora (UFJF), Juiz de Fora, MG, 36036-900 Brazil.

II. RELATED WORK

The continuous evolution of information and communication technologies has triggered a profound, multidimensional reconfiguration of the global audiovisual ecosystem. Recent literature points to a "pendulum swing" in the streaming market, characterized by the gradual and unmistakable reintroduction of operational strategies, monetization models, and programming structures historically associated with commercial broadcast television. Driven by severe saturation in mature consumer markets, stagnating subscriber acquisition rates, rising churn, and the skyrocketing cost of licensing major studio libraries, leading streaming providers have been forced to revise their initial narratives of differentiation and technological disruption. This trend of asymmetric convergence manifests in the introduction of hybrid, ad-supported monetization models (often in entry-level tiers lacking ad-skipping capabilities), the incorporation of live sports and entertainment broadcasts, the shift away from full-season "binge" releases in favor of sequential weekly episode rollouts, and the practice of service bundling. Thus, contemporary dynamics reveal a hybridization process in which on-demand services adopt the editorial, financial, and continuous-flow logic of broadcasting to ensure audience retention and the economic sustainability of their operations [6].

Against this backdrop of structural convergence between distribution models, TV 3.0 emerges as the primary driver of strategic differentiation and technological innovation for free-to-air television, redefining its functional relationship with the audience by delivering deep, contextual, and viewer-centric interactivity [2]. While on-demand digital services increasingly replicate the operational bottlenecks of traditional commercial TV, such as the mandatory display of unsolicited, uninterrupted advertising, the native infrastructure of TV 3.0 leverages the stability and predictability inherent in its editorial programming schedule to apply advanced Artificial Intelligence (AI) models, knowledge ontologies, and dynamic interface orchestration algorithms. This architecture enables the real-time, adaptive generation of graphic layouts and information layers that are deeply aligned with the active viewer's profile. By harmonizing the mass-reach efficiency of broadcast with the individualization capabilities of broadband, TV 3.0 transforms television viewing from a passive, uniform act into an immersive, participatory, and asymmetric experience. Thus, personalized interactivity establishes itself as a unique competitive advantage, restoring the prominence and appeal of free-to-air television in the face of the growing standardization of streaming catalogs.

Making this real-time adaptive delivery possible has intensified scientific efforts to develop methods for the automated semantic understanding of audiovisual content. Given that television programming is characterized by a marked heterogeneity of genres (such as news broadcasts, live sports, drama series, and entertainment shows), the automatic extraction of processable metadata to support personalization mechanisms presents complex analytical challenges. In this field, Wong et al. [3] report significant advances in the computational processing of live video streams through the application of AI

models focused on real-time detection, object segmentation, and motion prediction. Complementarily, Dorcheh et al. [4] emphasize the crucial importance of semantically identifying entities and events within media assets, demonstrating how the computational structuring of these elements maximizes the effectiveness of recommendation engines and content discovery systems. In both approaches, AI plays the role of converting raw computational signals into high-level semantic representations capable of powering intelligent decision-making systems.

Beyond the realm of media indexing and semantic analysis, recent academic research has focused on evaluating the impact of AI on user experience modeling and personalization. As demonstrated by Drofa [5], the application of machine learning algorithms for the dynamic adaptation of interfaces minimizes viewer cognitive load, optimizes the relevance of on-screen data, and simplifies the consumption of digital services across various contextual scenarios. These findings underscore AI's potential not merely as an engine for analytical data processing, but as an enabler of interactive environments that are adaptive and responsive to the context of use.

Despite these individual advances, the existing scientific literature treats semantic content analysis, user context profiling models, and interface orchestration mechanisms as predominantly isolated and disjointed research fields. While a significant portion of the work focuses on digital signal processing for metadata extraction, and another is dedicated to developing collaborative filtering engines for closed VoD ecosystems, a conceptual and practical gap exists regarding the development of integrated architectures. In particular, there is a scarcity of solutions capable of dynamically orchestrating the semantic understanding of linear streams, viewer context at runtime, and the adaptive rendering of interactive applications without compromising the broadcast flow. Furthermore, the operational constraints and prerogatives characterizing broadcasting, such as broadcaster governance, the preservation of editorial control, low-latency guarantees, and compliance with content rating guidelines, are often overlooked by conventional web-based personalization models.

Consequently, there is a clear need to design modular, extensible, and reusable architectures capable of bridging the gap between AI-driven semantic inference and dynamic interface orchestration, while upholding regulatory requirements and the editorial rigor of broadcasting. This work addresses precisely this gap in the literature by proposing and evaluating a unified architectural framework for semantic and context-aware personalization of interactive applications within the TV 3.0 ecosystem.

III. USE-CASE CHARACTERIZATION: A PREREQUISITE FOR ARCHITECTURAL DESIGN

Before detailing the proposed architecture, this section characterizes the use cases it is designed to support. This ordering is deliberate rather than incidental: as described in Section IV, the architecture's presentation layer is built around a Use-Case Templates Repository, whose components are ontological abstractions of the recurring interaction patterns

discussed below. Since these template components are derived directly from the use cases, the architectural design cannot be meaningfully introduced before the use-case vocabulary and requirements that shape it are established. The following analysis therefore precedes the architectural description, providing the conceptual foundation on which the template-based presentation layer is built.

The operationalization of interactivity in modern digital television relies on the predictability of broadcasters' editorial planning, where schedules and content agendas provide primary contextual metadata. The construction of the proposed architecture is guided by the formal abstraction of use cases, a requirement that dictates the creation, parameterization, and subsequent instantiation of visual templates within the presentation layer. Without this rigorous analysis, the graphical interface becomes rigid and unable to respond fluidly to the dynamic nature of audiovisual streams.

Mapping potential interactivity within programming reveals a functional redundancy that allows for moving beyond the simplistic link between an interface and a specific program. Consequently, an approach based on ontologies and the logical grouping of use cases is applied. This process establishes classes, properties, and relationships that enable the creation of a reusable, parameterized library of interface components.

This ontological categorization yields direct benefits for Artificial Intelligence processing. Restricting the catalog to well-defined components reduces ambiguity in the neural networks' decision space and simplifies the models' classification matrix. This optimization results in reduced inference latency, lower error rates, and minimized rendering friction throughout the pipeline, ensuring deterministic execution that meets the real-time demands of broadcasting.

Beyond computational optimization, ontological grouping enables the adaptive positioning of components. The graphical interface spatially adjusts across the screen, allowing for the coexistence of multiple simultaneous interactive elements, such as maintaining an audience chat in a peripheral area while triggering on-demand transactional modules (t-commerce) in other zones.

From a personalization perspective, decoupling the linear media stream from the interactivity layers allows a single broadcast transmission to generate diverse and asymmetric experiences across different receivers. This flexibility manifests in two interdependent dimensions: content asymmetry and application asymmetry.

In content asymmetry, the template structure remains constant, but the level of data detail is dynamically calibrated, allowing, for instance, an analytical viewer to see advanced sports telemetry while another receives simplified graphics. In application asymmetry, the system decides whether or not to instantiate a specific feature over the video; it might display a shopping channel for one profile and only informational data for another.

By harmonizing mass broadcasting with the conditional instantiation of elements via broadband, the system enables multi-level personalization. Coordination of this ecosystem is handled by the primary template (or "master template"), a supervisory framework that manages the user session, me-

diates profile-switching policies, and orchestrates the spatial rendering of secondary applications.

To underpin the ontological modeling and decision space of the Artificial Intelligence mechanisms, functional interactivity requirements were structured at a macro level based on the intended nature of engagement. This categorization guides the creation of a catalog of standardized, reusable templates, reducing ambiguity in the AI's decision space by defining how metadata extracted from the broadcast is visually translated within the interface.

A. Contextual and Informational Interactivity

The delivery of enriched information relies on templates designed to overlay data layers onto the broadcast without interrupting the primary audiovisual stream. The AI performs multimodal signal analysis to extract entities and events in real time, automatically populating interface fields with structured data from external APIs and semantic repositories.

In scenarios involving the tracking of sports metrics, the AI model identifies the context of the play and instantiates templates dedicated to displaying heat maps, athlete or team statistics, updated standings, and match schedules. In musical performances, AI correlates live audio with semantic databases to display templates featuring biographical data and up-to-date discographies.

In television news, language models and semantic analysis extract the topic under discussion to populate templates that generate QR codes or deep links sent to the user's smartphone. Similarly, in entertainment programs and reality shows, the system identifies participants who are on screen or mentioned in real-time, populating templates that allow viewers to access a participant's history, rewatch key moments, or view biographical details.

B. Manipulation and Control of the Audiovisual Flow

Management of the audiovisual stream relies on media orchestration templates that grant the viewer autonomy over the visual perspective and the content's timing. AI analyzes the multiple video signals available within the broadcast infrastructure and populates an interactive mosaic-style template, allowing the viewer to seamlessly switch camera angles and the corresponding audio output.

Computational models also manage Picture-in-Picture (PiP) windows to automatically position secondary feeds, such as replays of key moments, without obscuring the main event. Regarding the temporal dimension, AI applies segmentation algorithms to identify key events and populate timeline templates with semantic markers (bookmarks). This automation creates dynamic shortcuts to sports plays, musical segments, or thematic news blocks.

Based on these markers, AI configures "catch-up" templates for the on-demand display of previous episodes, behind-the-scenes footage, and supplementary narratives (spin-offs). Additionally, the architecture enables the generation of templates featuring vertical video clips (in a social media format) containing the program's highlights. Viewers can watch these clips directly on the screen or access them via QR code for instant sharing on their personal social networks.

C. Contextual Commerce (T-Commerce)

In this scenario, computer vision algorithms perform object detection, segmentation, and tracking within the scene, ontologically linking objects to product catalogs to instantiate and populate templates in real time.

This integration allows AI to identify clothing, costumes, sports equipment, or decor visible in the video, populating the interface with descriptions, prices, and purchase buttons. AI also manages graphic components for selling official merchandise, ticket sales for future events, and displaying customized sponsor offers tailored to the viewer's profile.

In broadcasts featuring tutorials or sequential procedures (such as cooking or home improvement shows), AI extracts the displayed ingredients and steps to populate an interactive guide template. This component allows users to browse the list of items, save the digital file or recipe to their smartphones, or purchase the necessary items directly through integrations with partner e-commerce platforms.

D. Engagement, Gamification, and Collective Participation

Finally, collective participation and engagement mechanisms utilize interactive templates designed to transform the viewing experience into something even more immersive. AI analyzes the broadcast timing and audience context to launch real-time voting and polling templates, processing audience votes and instantly updating on-screen graphics.

This dynamic drives everything from selecting the next song at a concert to steering narrative paths in dramatic productions. In fictional contexts, collective or individual decisions feed the AI's rule engine, triggering templates with personalized alternative outcomes and endings.

The gamified aspect is structured around quiz and trivia templates that the AI automatically generates and synchronizes with events detected in the video, calculating scores in real time. For social interactivity (Social TV), AI curates and automatically filters audience comments before populating the chat template; this template can remain fixed in a sidebar while other secondary components rotate on the screen.

IV. PROPOSED ARCHITECTURE

Figure 1 presents the proposed architecture for semantic and context-aware personalization in TV 3.0 environments. The architecture integrates semantic content analysis, recommendation mechanisms, viewer-context modeling, and dynamic application composition within a unified personalization architecture. Its modular organization promotes extensibility and facilitates the incorporation of new personalization use cases while preserving broadcaster control over content curation and service delivery. The architecture comprises four main components: the Use-Case Templates Repository, the Applications Repository, the Business Model, and the Database.

The architecture adopts a modular design that separates content analysis, recommendation, personalization, application composition, and data management functions into specialized components. This organization facilitates the evolution of individual modules, supports the incorporation of additional

personalization use cases, and promotes scalability, maintainability, and interoperability across the TV 3.0 ecosystem.

Personalization emerges from the interaction among semantic descriptors extracted from audiovisual content, viewer-context information, and use-case templates. The Business Model component acts as the architectural core, coordinating the exchange of information among repositories, data sources, and personalization services. This arrangement enables the dynamic selection and configuration of application elements according to both the characteristics of the content being presented and the contextual attributes of the active viewer.

The following subsections describe the responsibilities and interactions of each architectural component.

A. Use-Case Templates Repository

Designing personalized experiences in modern digital television environments relies on the premise that viewer engagement patterns often transcend specific program genres or rigid content categories. Foundational use cases—such as contextual TV commerce (t-commerce), real-time statistics and highlights, narrative summaries, catch-up services, support for sequential viewing (binge-watching), Social TV features, and the generation of personalized channels, prove to be cross-cutting and applicable to a wide range of audiovisual broadcasts. Consequently, interface personalization cannot be defined solely by pre-established program formats; it requires an approach that synergistically addresses intrinsic interaction goals and the contextual conditions inherent to each viewing session.

To enable this versatility, the Use Case Model Repository acts as the persistence layer responsible for storing pre-defined, ontologically modeled structures that formalize these recurring interaction scenarios. Each use case model rigorously specifies a set of graphical components, spatial presentation rules, navigation structures, and behavioral constraints associated with a specific interactive purpose. By serving as modular, highly reusable building blocks for the continuous composition of adaptive experiences, these models ensure interface flexibility without compromising the visual identity and editorial guidelines established by the broadcaster.

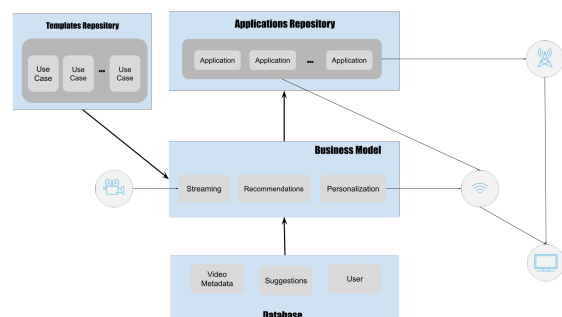


Fig. 1. Architecture overview

The primary template serves as a control repository for the broadcaster's full catalog of interactive features. This functional centralization formally resolves the tension between the system's predictive automation and user agency: while the architecture delivers fluid experiences automated by Artificial Intelligence algorithms based on the current programming schedule, it strictly safeguards viewer autonomy. Consequently, the user retains the freedom to access informational data, retrieve past interactive events, or manually launch applications that were not algorithmically suggested for their profile. The primary template thus consolidates a harmonious convergence between the system's predictive mediation and the viewer's deliberate choice.

B. Applications Repository

The Application Repository serves as the persistence and lifecycle management layer for custom software instances. Far from being a mere static file store, this component operates as a dynamic, indexed repository that centralizes the interactivity artifacts resulting from the fusion of the interface abstraction layer and the decision-making engine. These instances are generated through a systematic process of selecting, parameterizing, and instantiating use-case models and ontological templates. This process is strictly guided by a synthesis of semantic content descriptors—characterizing the ongoing audiovisual stream—alongside user context variables (such as profile, preferences, and device capabilities) and strategic guidelines derived from the Business Model component.

The convergence of these variables within the governance layer results in a precise specification of the software artifact to be instantiated and made available by the repository. Consequently, while the underlying linear audiovisual stream is transmitted in an agnostic manner, it unfolds into multiple application configurations at the final presentation layer. This service-oriented approach enables a heterogeneous and asymmetric viewing experience, allowing different receivers to display a range of content, from subtle informational and statistical overlays to transactional graphic elements, hyper-personalized recommendations, and contextual services tailored to specific consumption profiles. In short, the Application Repository facilitates the transition from the traditional mass-broadcast model to an adaptive, user-centric ecosystem, ensuring low-latency artifact retrieval and the integrity of the personalized experience delivery.

C. Business Model

The Business Model component constitutes the architectural core and coordinates the processes responsible for semantic content analysis, recommendation generation, and viewer-context modeling. It comprises three integrated modules:

1) *Streaming*: This module is responsible for the continuous ingestion and analysis of real-time audiovisual streams, employing artificial intelligence techniques to automatically extract high-level semantic descriptors. Through multimodal signal segmentation, the component identifies and categorizes structural content elements, such as visual entities, settings,

narrative events, dialogue transcripts, and contextual keywords. The formal semantic representation derived from this processing establishes a structured metadata base that directly feeds the inference engines and personalization algorithms in the architecture's subsequent layers.

2) *Recommendations*: This module acts as the architecture's inference core, responsible for selecting, combining, and parameterizing use-case models based on the alignment between content semantic descriptors, the user's contextual state, and operational personalization strategies. It ensures the generation of coherent visual and functional arrangements, even in scenarios characterized by uncertainty, data sparsity, or a complete lack of viewer profile information.

3) *Personalization*: This module is responsible for instantiating, maintaining, and continuously updating the user context model by synthesizing data derived from audiovisual consumption history, interface interaction patterns, explicitly stated preferences, and contextual metadata provided by the TV 3.0 middleware layer. By providing a structured abstraction of the viewer profile, the component underpins and guides the decision-making of personalization algorithms in the architecture's upper layers. Concurrently, the module incorporates rigorous mechanisms to ensure that the collection, processing, and retention of sensitive viewer data strictly comply with data protection and privacy regulatory frameworks, notably the General Data Protection Law (LGPD).

D. Database

The architecture maintains three specialized data repositories that support semantic analysis, recommendation generation, and viewer-context management.

1) *Video Metadata*: Stores semantic descriptors extracted from audiovisual content, providing the content representation used by the personalization processes.

2) *Suggestions*: Stores use-case templates, recommendation rules, and personalization knowledge used by the recommendation module during experience composition.

3) *User*: Stores viewer-context information, including viewing history, preferences, and active profiles, managed by the personalization module.

E. Interface Delivery Mechanisms

The architecture supports two complementary delivery mechanisms for personalized applications.

1) *Over-the-Air*: Intended for receivers without broadband connectivity or unavailable viewer-context information. In this scenario, generic application instances are delivered using locally available content and interaction resources.

2) *Via the Internet*: Activated when broadband connectivity is available, enabling access to viewer-context information and full execution of the personalization architecture. Under these conditions, applications can be dynamically instantiated and adapted according to semantic descriptors, viewer context, and selected use-case templates.

V. PERSONALIZATION PROCESS

The proposed architecture establishes personalization as a stochastic, multi-stage decision-making process that transcends the mere recommendation of audiovisual content. Rather than limiting adaptation to the displayed informational data, the system determines, predictively and analytically, which use-case models should be activated, how they should be combined into spatial arrangements, and how the resulting application should be instantiated for a specific viewing session.

This operational flow begins with signal ingestion and multimodal analysis, from which semantic content descriptors are extracted and viewer context variables are collected (when available). The Business Model component synthesizes these input vectors to map the personalization objectives best suited to the current scenario. To support this inference, the persistence layer stores and retrieves semantic descriptors, knowledge graphs, ontological use-case models, recommendation rules, and user profiling data. Based on this convergence, the recommendation module selects, orchestrates, and parameterizes one or more use-case models capable of supporting the intended experience.

Use-case selection relies on the premise that viewer engagement patterns are cross-cutting and applicable to multiple program genres. Recurring scenarios, such as contextual t-commerce, real-time statistics and highlights, narrative summaries, catch-up services, support for sequential consumption, and Social TV, are structured as modular, reusable building blocks. The ontological grouping of these use cases reduces the redundant proliferation of templates and optimizes the search space for Artificial Intelligence algorithms; this lowers inference latency and interface selection error rates, while enabling the simultaneous, harmonious execution of multiple on-screen applications (such as overlaying a chat channel with a transactional stream).

Once selected, use-case models are instantiated and dynamically populated with data, recommendations, interactive graphical elements, and contextual services. As a result of this orchestration, viewers exposed to the same linear broadcast stream may receive heterogeneous and asymmetric application configurations, tailored to their preferences, interaction history, and device context. Personalization thus unfolds across three interdependent levels: the ontological selection of use cases, the spatial and structural composition of the application, and the adaptation of the presented content.

The resulting application instances are consolidated in the Application Repository, which serves as the integration and lifecycle management point for the generated software artifacts. This component ensures low-latency retrieval of the interactive artifacts that will ultimately be rendered on the receiver's presentation layer.

This entire visual and functional orchestration is governed by the Master Template, a non-intrusive supervisory framework equipped with a dynamic visibility protocol. Designed to preserve immersion in the main audiovisual stream, the Master Template operates in the background and automatically hides itself after brief periods of user inactivity or during

the spontaneous execution of context-triggered interactive elements, reappearing only upon the viewer's explicit command. In addition to managing profiles and spatial transition policies, the Master Template centralizes the broadcaster's complete catalog of interactive features. This functionality formally resolves the dilemma between system predictive automation and user autonomy: while AI delivers adaptive experiences fluidly and automatically, the viewer retains the freedom to browse the catalog to retrieve past interactions, consult informational data, or manually launch applications not directly suggested for their profile.

The use-case model-driven approach leverages the nature of television production, wherein predictable editorial formats and engagement opportunities are translated into highly reusable personalization patterns. This feature allows the architecture to anticipate interaction scenarios at design time while maintaining the flexibility required for dynamic reconfigurations at runtime.

The interface's degree of adaptability depends on the density of available information. In purely broadcast scenarios, personalization relies on the media stream's semantic descriptors and editorial rules predefined by the broadcaster. Given the availability of broadband connectivity and profiling data, the architecture progressively enriches the experience through the dynamic selection, composition, and instantiation of multiple use cases. This mechanism ensures graceful degradation, guaranteeing that all receivers a basic functional layer of interactivity, while enabling increasingly personalized experiences as new contextual variables are assimilated.

Finally, the same inference and orchestration mechanisms apply with equal effectiveness to Video on Demand (VOD) environments. Interactive elements, recommendations, and contextual services adapt agnostically to both continuous streaming and asynchronous consumption. As a result, the architecture establishes a unified personalization model for linear and on-demand television, enabling AI-driven adaptive experiences across the entire TV 3.0 ecosystem.

VI. PRELIMINARY EVALUATION

The proposed architecture was qualitatively evaluated regarding its ability to support semantic and context-aware personalization within TV 3.0 ecosystems. A key finding concerns the strict decoupling of personalization logic from application structure, enabled by the use of reusable, ontologically structured use-case models. This approach allows recurring, cross-cutting engagement patterns, such as contextual t-commerce, real-time statistics, Social TV, and support for sequential consumption, to be represented independently of specific programs or genres. Consequently, system scalability is enhanced, and new use cases can be continuously incorporated without requiring restructuring of the underlying software architecture.

A second aspect relates to the direct computational performance gains achieved through the ontological grouping of these use cases. By mitigating the redundant proliferation of similar functional templates, the architecture constrains the search space for Artificial Intelligence algorithms. This results in a measurable reduction in model inference latency and

minimizes error rates in template selection, ensuring the fluid, automated population of interfaces with minimal processing overhead.

From a user experience perspective, the adoption of a multi-level personalization model stands out. Beyond merely adapting informational content and recommendations, the architecture orchestrates the selection, spatial composition, and dynamic instantiation of use-case models. Consequently, the same linear audiovisual stream unfolds into deeply heterogeneous and asymmetric application configurations, varying the level of data depth, interactive graphic components, and interaction modalities based on content characteristics and the active viewer's profile.

Governance of this adaptive delivery is ensured by the primary template, whose role has proven crucial for usability and visual non-intrusiveness. Its dynamic visibility protocol preserves immersion in the broadcast by automatically hiding during periods of inactivity, while its central on-demand catalog function resolves the tension between automation and agency: although AI delivers predictive, context-driven interactive features, the viewer retains the autonomy to retrieve past events or actively access applications not automatically suggested for their profile.

Analysis further indicates that the architecture supports progressive personalization through graceful degradation, adjusting to the density of available information at runtime. In purely broadcast scenarios, personalization relies on semantic signal analysis and the broadcaster's editorial rules. When broadband connectivity and user profiling data are available, the system enriches the experience by dynamically composing multiple use cases, operating agnostically across both linear streams and Video on Demand (VOD) services.

Finally, the proposed architecture integrates, within a unified conceptual and operational framework, AI-based semantic content analysis, continuous modeling of viewer context in compliance with LGPD principles, and dynamic application composition managed by a governance template. This systematic integration addresses a critical gap identified in the literature, where these layers are often treated as isolated and disjointed processes. Preliminary results confirm the technical feasibility of the solution and its high potential to underpin ecosystems of personalized and adaptive interactivity in TV 3.0.

VII. FINAL REMARKS

This work presented an extensible and modular architecture aimed at semantic and context-aware personalization within TV 3.0 ecosystems. The proposed approach unifies—within an integrated conceptual and operational framework—the semantic analysis of media streams, dynamic modeling of viewer context, governance of use-case ontological models, and adaptive application composition. This integration demonstrates the capability for agnostic and resilient operation across both traditional broadcast environments and broadband networks.

Preliminary evaluations confirm that the architecture supports personalization across multiple interdependent levels, encompassing the logical selection of use cases, the spatial

orchestration of interfaces, and the fine-grained adaptation of displayed content. By correlating semantic descriptors extracted from the audiovisual signal with viewer variables, the system enables the seamless generation of asymmetric interactive experiences in both linear services and Video on Demand (VOD).

The scientific contribution of this work lies in advancing the television personalization paradigm beyond traditional recommendation engines by introducing a flexible framework based on reusable components. This infrastructure empowers broadcasters to deliver heterogeneous interactive scenarios while strictly maintaining editorial control, computational scalability, and platform extensibility.

Future work will focus on developing a functional prototype and experimentally validating the architecture in real-world TV 3.0 broadcasting scenarios, thereby empirically verifying the effectiveness of multi-level personalization. Separate from this validation, future work will incorporate the modeling of operational and logistical constraints that influence the system's real-time recommendations. In the context of e-commerce, the research will investigate the incorporation of dynamic physical logistics constraints (such as inventory availability, delivery times, and shipping routes) into the recommendation engine, thereby preventing inefficiencies in product conversion and delivery. Conversely, for the provision of digital content and services, the system will apply analogous constraints focused on data orchestration and delivery, subjecting the display of interfaces and the delivery of interactive flows to the current capacity, latency, and operational costs of the scalable network infrastructure (broadcast and broadband).

At the same time, with regard to improving the mechanisms for selecting and combining use-case models, the aim will be to focus on training more robust machine learning algorithms, incorporating techniques to defend against adversarial attacks and rigorous metadata validation to mitigate risks of data poisoning or inferential bias in the rendering of interfaces. Regarding dynamic instantiation and runtime management of the Application Repository, the goal will be to implement confidential computing routines and strict data flow controls, ensuring that the assembly and delivery of interactive artifacts preserve the viewer's privacy without exposing their telemetry to vulnerabilities. Finally, the evaluation of the user experience in live and on-demand services will be accompanied by the integration of principles of explainable AI and transparency into the master template, allowing the public to understand the recommendation criteria and actively manage their privacy preferences, in strict compliance with the General Data Protection Law (LGPD).

REFERENCES

- [1] J. M. I. P. Josué, P. A. Valentim, and D. C. M. Saade, "TV 3.0: Definição e Uso de Perfil de Telespectador no Ambiente de TV Digital Aberta," In: *Workshop Futuro da TV Digital Interativa - Simpósio Brasileiro de Sistemas Multimídia e Web (WebMedia)*, 29., 2023, Ribeirão Preto/SP. Anais [...]. Porto Alegre: Sociedade Brasileira de Computação, 2023, pp. 171–177, doi: https://doi.org/10.5753/webmedia_estendido.2023.236156.
- [2] M. F. Moreno *et al.*, "R&D Progress on TV 3.0 Application Coding Layer," *SET International Journal of Broadcast Engineering*, 2023, doi: <https://doi.org/10.18580/setijbe.2023.1>.

- [3] G. Wong, Y. Kumar, J. Li, and D. Kruger, "Real-Time Object Detection and Skeletonization for Motion Prediction in Video Streaming," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 39, no. 28, pp. 29712–29714, 2025, doi: <https://doi.org/10.1609/aaai.v39i28.35376>.
- [4] D. S. M. Dorcheh *et al.*, "SmartCrop: AI-Based Cropping of Soccer Videos," in *IEEE International Symposium on Multimedia (ISM)*, Laguna Hills, CA, USA, 2023, pp. 20–27, doi: <https://doi.org/10.1109/ISM59092.2023.00009>.
- [5] D. Drofa, "Integrating Advanced API Solutions into Full-Stack Web and Mobile Applications to Optimise User Experience," *International Journal of Current Science Research and Review*, vol. 8, no. 5, 2025, doi: <https://doi.org/10.47191/ijcsrr/V8-i5-16>.
- [6] M. Meimaridis and B. Campanella, "'A TV voltando a ser TV': a convergência do streaming de vídeo com práticas televisivas," *Revista FAMECOS*, vol. 32, no. 1, p. e46764, 2025, doi: <https://doi.org/10.15448/1980-3729.2025.1.46764>.
- [7] J. dos Santos, L.-C. S. C. S. Sousa, C. S. Neto, G. Lemos, S. Colcher, M. F. Moreno, and D. Muchaluat-Saade, "An Overview of the TV 3.0 Application Coding Layer Specification," *IEEE Transactions on Broadcasting*, vol. 72, no. 2, pp. 627–641, 2026, doi: <https://doi.org/10.1109/TBC.2026.3689351>.
- [8] Fórum SBTVD, "O que é a TV 3.0, como funciona e o que muda nos canais abertos do Brasil," *Fórum do Sistema Brasileiro de Televisão Digital Terrestre*, 2024. [Online]. Available: <https://forumsbtvd.org.br/dtv-o-que-e-a-tv-3-0-como-funciona-e-o-que-muda-nos-canais-abertos-do-brasil/>.



Letícia Concolato Rangel holds a graduate degree in Electrical Engineering from Estácio de Sá University and is a master's student in the Graduate Program in Computer Science (PPGCC) at the Federal University of Juiz de Fora (UFJF). She has experience in the fields of Computer Science and Telecommunications Engineering, working within the TV 3.0 ecosystem, with an emphasis on the orchestration of video transmission and reception flows, the development of interactive applications, and content personalization based on Artificial In-

telligence. Her main research interests include multimedia systems, adaptive frameworks, and interactive digital television.



Marcelo F. Moreno is an Associate Professor at the Federal University of Juiz de Fora (UFJF), Brazil. He holds a Ph.D. in Computer Science from PUC-Rio and works in the areas of multimedia systems and computer networks. From 2022 to 2023, he was a Visiting Professor at the International Audio Laboratories Erlangen (FAU/Fraunhofer IIS), Germany. He co-edited several ITU-T and ITU-R Recommendations and Technical Reports, coordinating ITU-T working groups for over a decade. Since 2015, he has led the Application Coding Working Group of

the Brazilian Digital TV System Forum (SBTVD) and serves as editor of the ABNT standards for TV 2.5 and TV 3.0 on middleware and application coding. His research focuses on application-oriented broadcasting, second-screen integration, audience measurement, and privacy-aware media platforms for next-generation digital television. He has published widely in the field and is a member of the Brazilian Society of Television Engineering (SET), the Institute of Electrical and Electronics Engineers (IEEE) and the Brazilian Computer Society (SBC).