Data Processing and Analysis to Improve the Management of Broadcasting Grants in Brazil

Higo T. P. Silva, Pedro H. D. Almeida, José A. S. Torres, Uanderson A. P. Frota, Pedro D. Alvim, Daniel P. Gonçalves, Hugerles S. Silva, Ugo S. Dias, Daniel G. Silva, Thiago A. Soares and Thiago R. Silva

***Abstract*—The Brazilian Ministry of Communications (MCom) currently utilizes the Mosaico platform for managing broadcasting grants. However, this platform primarily focuses on channel data, which does not align directly with the requirements of grant management. Consequently, there arises an imperative need for the development of a dedicated grants management system (GMS). To address this need effectively, it is crucial to undertake a thorough sanitation process for channel data. The objective is to identify errors and inconsistencies within the records and to adapt this data to a structure tailored to broadcasting grants. In this context, this paper presents the results of data processing derived from a python module named Mosaico database analysis (MDbA). The MDbA serves the purpose of automating the detection of errors and inconsistencies within the Mosaico channel database. It achieves this by applying predefined processing rules, which were defined in conjunction with a MCom team. Additionally, the module facilitates the creation of a new data structure specifically designed for grants, referred to as the grant data structure (GDS). As a result of the analysis conducted by MDbA, approximately 46941 inconsistencies of various types were identified within the Mosaico data. Moreover, leveraging a novel methodology for aggregating channel data in grants, approximately 23257 GDSs were established.**

***Index Terms*—Broadcasting, data processing, grants.**

I. Introduction

THROUGHOUT history, the Federal Executive Branch has held the responsibility for attributing, renewing, and managing broadcasting concessions in Brazil [1]. This practice was formalized by Decree 20.047 on May 27, 1931, which delegated the regulation of radio broadcasting activities to the Federal Government [1]. Based on the trusteeship model outlined in the 1934 Constitution, the state assumed the duty of meticulously managing the electromagnetic spectrum. Additionally, the 1934 Constitution introduced a command-and-control model, requiring a license to be obtained for the use of the electromagnetic spectrum, delineating its terms of usage [2], [3].

Substantial alterations in the administration of broadcasting licenses transpired in 1967 with the enactment of Decree-Laws 236 and 200 (February 25) [1]. Decree-Law 236 of 1967 imposed numerical restrictions on the ownership of various types of broadcasting licenses and introduced a new educational broadcasting license category [4]. Simultaneously, Decree-Law 200 established the Ministry of Communications (MCom), which is the government body responsible for managing and renewing broadcasting licenses [5]. Another pivotal entity in Brazilian broadcasting emerged in 1997 with the establishment of the National Telecommunications Agency (Anatel) [6]. As stipulated by the General Telecommunications Law (Law 9.472 of July 1997), the MCom retains responsibility for managing broadcasting grants, while Anatel is mandated to regulate and supervise the planning and utilization of the electromagnetic spectrum from a technical engineering perspective [1], [2]. This management framework, whereby the MCom handles the licensing procedures and Anatel oversees the technical control of electromagnetic channels, remains in effect to the present day.

The division of tasks assigned to MCom and Anatel creates difficulties for a unified management of Brazilian broadcasting licenses. Currently, MCom depends on data from the Mosaico platform for managing broadcasting licenses [7]. This platform, however, was originally developed by Anatel to support technical management and oversight of broadcasting channels, providing specific information on channel characteristics rather than licensing details. Consequently, there is a misalignment between the grant-related data MCom needs and the channel-specific information that Mosaico offers. This discrepancy introduces technical obstacles that undermine MCom's efficiency in licensing management. In the absence of a dedicated system for grant management, MCom resorts to labor-intensive processes, manually converting and adapting data from Mosaico to serve its licensing needs. From a grant management perspective, an optimal solution would entail a platform structured to accommodate grant-related data rather than channel specifics.

In contrast to the Brazilian approach, in the United States, the Federal Communications Commission (FCC) manages broadcasting licenses through the licensing and management system (LMS), which integrates technical channel data with licensing information in a unified platform [8], [9]. Similarly,

H. T. P. Silva (higo.silva@ee.ufcg.edu.br, ORCID 0000-0002-3680-9659) is with the Electrical Engineering Department, Federal Institute of Paraíba (IFPB), Paraíba, Brazil.

P. H. D. Almeida (phdornelas.almeida@gmail.com, ORCID 0000-0002-7187-9986), P. D. Alvim (pedro.alvim@redes.unb.br, ORCID 0009-0003-9472-5300), H. S. Silva (hugerles.silva@unb.br, ORCID 0000-0003-0165-5853), U. S. Dias (udias@unb.br, ORCID 0000-0003-0121-1523) and D. G. Silva (danielgs@unb.br, ORCID 0000-0003-2858-1110) are with the Electrical Engineering Department, University of Brasília, Brasília, Brazil.

J. A. S. Torres (alberto.torres@redes.unb.br), U. A. da P. Frota (uanderson.aguiar@gmail.com) and D. P. Gonçalves (dani.pereira47@gmail.com, ORCID 0009-0000-0480-4936), are with the LATITUDE Laboratory, University of Brasília, Brasília, Brazil.

T. A. Soares (thiago.soares@mcom.gov.br) and T. R. Silva (thiago.silva@mcom.gov.br) are with the Brazilian Ministry of Communications (MCom), Brasília, Brazil.

This work has been funded by TED UnB-MCom 01/SERAD/2022.

Submission: 2024-06-04, First decision: 2024-11-24, Acceptance: 2024-12-01, Publication: 2024-12-04

Digital Object Identifier: 10.14209/jcis.2024.22

in Canada, the Canadian Radio-Television and Telecommunications Commission (CRTC) administers broadcasting licenses through a dedicated system [10]. Since each country has its own specific regulation regarding broadcasting services, Brazil cannot adopt a foreign license management system because it would not meet the requirements of the unique legal framework that governs national communications.

In this context, a collaborative development initiative was started between the MCom and the University of Brasília (UnB) to prototype a platform aimed at enhancing agility and transparency in the management of broadcasting grants in Brazil. This new platform is currently in the development phase and, with the direct involvement of a specialized MCom team, is being designed to meet the technical requirements for streamlining the management of broadcasting grants. For synthesis purposes, the system under development is generically referred to throughout this work as grant management system (GMS), as it does not yet have an official name for disclosure.

A crucial phase in the development of the GMS involves transforming the channel-specific data stored in the Mosaico database into grant-related information. Moreover, conducting an analysis to identify any discrepancies is essential to ensure the robustness and efficiency of the ongoing system in its management capabilities. Firstly, it is necessary the automatic identification of information inconsistencies in the Mosaico database under the aegis of Brazil's communications legislation. Hence, this work originally proposes the development of a tool named Mosaico database analysis (MDbA). It is a Python-based module which addresses the objectives of processing, identifying inconsistencies, and converting brazilian broadcasting channels data into the associated grant data.

The MDbA module operates on two primary data processing fronts. Firstly, it conducts a comprehensive analysis to identify inconsistencies and registration errors within the Mosaico channel database. In context, Anatel relied on a system lacking data curation functionalities until the Mosaico system was introduced in 2016, significantly improving data quality. The entire process of evaluating and validating information was manual with the previous legacy system. Due to this change in systems, only new records enjoy enhanced reliability. Thus, it is imperative to implement processing, validation, and sanitation procedures for unstructured broadcast data, ensuring that future data consumption by the GMS is based on clean data.

Secondly, the MDbA is tasked with adapting channel-specific information into grant information. This transformation involves aggregating data from various broadcast channels into a unified grant record, complete with a unique identifier. Building upon this aggregation, critical information for the ministry is derived, including the grant's reference date, expiration date, and the count of renewal cycles. It's crucial to note that presently, this information is computed manually, significantly hindering the efficiency of grant management by MCom. Furthermore, the MDbA undertakes another vital process concerning the constraints outlined in Decree-Law 236. This process is indispensable for delineating any potential irregularities in the boundaries of properties granted by individuals or legal entities. Therefore, the process of unifying data into a new structure is a necessity for the future GMS, given that a new model is needed to meet the requirements of grant management. Furthermore, it would not be possible to evaluate rules and perform corrections without a unified data structure. It is important to emphasize that MDbA was developed specifically to process and generate data related to broadcasting licenses, such as television and radio services. Therefore, this module is not applicable to other business models, such as cellular and satellite communications, due to the inherent differences in data and legal framework.

The rest of this work has the following structure. Section II outlines the developmental stages of MDbA, beginning with an overview of channel data in Subsection II-A, followed by a detailed description of the validation rules implemented in MDbA in Subsection II-B. The process for generating grant data is expounded upon in Subsection II-C, followed by a explanation of the limits imposed by Decree 236, in Subsection II-D. Section III presents the numerical outcomes derived from the research. Finally, in Section IV, the conclusions of the work are exposed.

II. Data Processing

The process of collect and analyse data have become crucial for governments to make informed decisions and improve the efficiency of their operations [11]. However, realizing the potential of government data requires a structured and well-defined approach. In order to meet the requirements stipulated by MCom for the development of the MDbA and the GMS, activities were carried out to comprehend both the business and Anatel's data, since channel information did not necessarily represent useful information about the grant itself. This is where the cross-industry standard process for data mining (CRISP-DM) model can be highly beneficial [12]. The CRISP-DM model provides a standardized method for data mining projects, making it particularly useful for processing the Mosaico's data. Government agencies can ensure that their data evaluation activities are conducted in a methodical and results-driven manner, ultimately leading to better governance and public service delivery, by following the systematic approach of CRISP-DM [13]. This approach has proven to be particularly valuable in handling complex data analysis tasks related to areas such as public health, urban planning, and resource allocation [14].

This model consists of six phases - business understanding (BU), data understanding (DU), data preparation (DP), modeling (MO), evaluation (EV), and deployment (DE) - that guide the entire data processing life cycle [15]. These phases are iterative, meaning that feedback from later phases may result in additional work in earlier phases. The first phase of the CRISP-DM model is the BU. During this phase, the focus is on gaining a deep understanding of the objectives, goals, and requirements in relation to the data processing project. The understanding of the business context is crucial in order to align the data processing activities with the overall strategic goals of the MCom [16]. During this phase, several meetings were conducted between the UnB and MCom teams to outline the requirements for the proposed GMS. The objective was to gain insight into the challenges encountered in managing broadcasting grants.

In turn, the DU phase is the second stage of the methodology. This phase emphasizes the exploration and analysis of the available data. This includes tasks such as collecting, reviewing, and analyzing the data to identify any patterns, inconsistencies, or missing values. Additionally, data profiling and data quality assessment are conducted during this phase to ensure that the data is suitable for further processing. Similar to the BU phase, this stage involved periodic meetings to familiarize teams with the channel data available on the Mosaico platform. In addition, data processing rules were established based on the nuances of Brazilian communications legislation, which are integral to the development of the GMS. The next stage, the DP phase, involves tasks such as data cleaning, transformation, integration, and formatting. During this phase, the UnB group worked closely with the MCom team to prepare the raw data for analysis. This included identifying inconsistencies or missing values, as well as merging different datasets and ensuring that the data was in the appropriate format for further analysis and modeling.

The MO step is the fourth phase in the CRISP-DM model. During this phase, various modeling techniques are applied to the prepared data in order to derive meaningful insights in the data that can inform decision-making and support the objectives of the MCom in improve the management of brazilian broadcasting. This phase involved the application of modeling and data processing techniques using the MDbA python module. The module encapsulated the implementation of the rules defined in the DU phase, facilitating the automation of error detection and inconsistency resolution within Mosaico records. Additionally, it encompassed the implementation of rules governing the creation of the new grant data structure. The last step is the EV, in which the CRISP-DM model focuses on assessing how well the developed models meet the business goals and criteria. During the EV phase, the UnB tem assessed whether the rules implemented in MO stage meet the predefined success criteria and if they are aligned with the business goals of the MCom. Joint evaluations between teams were conducted to validate the effectiveness and suitability of the outcomes.

In this section, the MDbA module, developed primarily in the DU and DP phases, is described in detail. The MDbA Python module serves two primary objectives. Firstly, it aims to identify inconsistencies and errors within records stored on the Mosaico platform. Leveraging data intelligence gathered during the DU phase, MDbA incorporates multiple rules to automate error detection. This automation empowers the MCom team to concentrate their efforts on correcting the database, ensuring the quality of data consumed by the ongoing GMS. A methodology similar to the hypertext transfer protocol (HTTP) standard return codes is employed, wherein each numerical code is related to a specific error. The second application of MDbA's data processing involves generating instances of grant data. This process encompasses aggregating and intersecting disparate information from associated channels to define a structure uniquely characterized by a specific grant identifier (GID). Details regarding broadcasting licenses are extracted from this aggregation process, including reference and validity dates, as well as the number of renewal cycles associated with each license.

For better visualization and understanding, Fig. 1 presents the MDbA process diagram. This diagram is structured according to the business process model and notation (BPMN) [17]. In this diagram, the first two activities, represented by the two upper blocks, are related to data verification and validation, where multiple rules are applied to certify data adequacy. These rules are explained in detail in Sections II-A and II-B. In turn, in the last task, represented by the lower block in Fig. 1, the processes of generating grant data, in addition to another validation phase according to Decree 236, are executed. These processes are explained in Sections II-C and II-D.

A. Channel Data Description

During the DU and DP phases, the MCom team provided a Mosaico's database dump, containing all the channel data available in the platform. This dataset facilitated an in-depth analysis of the Mosaico data structure and the development of error detection applications in the MDbA module. The main information contained in the channel data fields are described below. For a more detailed view, a direct access to the Mosaico page is recommended [7].

Each channel record is distinguished by a unique channel identifier (CID) code predefined by Anatel. Additionally, the telecommunications inspection fund (Fistel) code, established in Law 5.070 of July 7, 1966 [18], plays a crucial role in channel identification, granting the respective entity access to invoices, inspection fees, and fines associated with spectrum usage [19]. Mosaico also offers the basic information on license-holding entities, including the national register of legal entities (NRLE) and the corporate name of the corresponding entity. In order to check legal information about entities holding broadcasting licenses and their respective partners, the MDbA module incorporates information from Brazilian Federal Revenue, provided by the MCom team.

Anatel employs specific indicator codes to classify the procedural status of broadcasting channels. A comprehensive list of the status codes and their descriptions is detailed in Tab. I. It can be seen that each code indicates a specific situation in the life cycle of a broadcasting channel, which goes through several technical and legislative processes [1]. Throughout this work, a channel will be designated as inactive (unused) if it has a status code `C0` (vacant channel) or `CE-01` (deleted channel)[1]. Otherwise, the channel is denominated as active. Furthermore, Mosaico provides the service modality of a broadcasting channel. In this work, the following generation services are addressed:

- Analog television transmission (TV);
- Digital television transmission (GTVD);
- Sound transmission operating in medium waves with amplitude modulation (OM);

[1]The term "inactive" used in this work does not correspond to technical terminology. In this text, inactive channels are defined as those not currently in use, in this case, channels with a status of `C0` or `CE-01`. However, in the technical terminology, channels with a status in `C5` are also categorized as inactive, as they have pending granting.

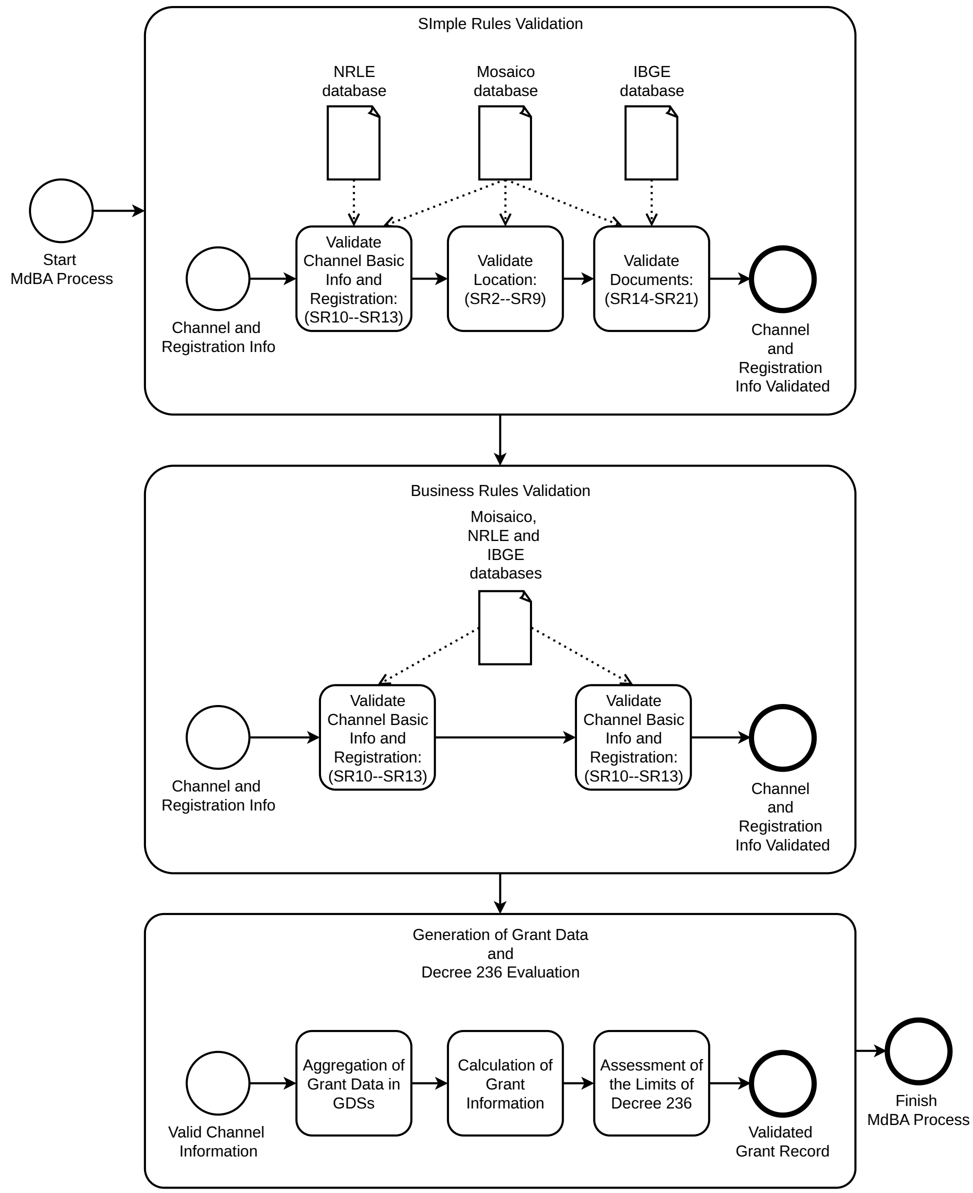


Fig. 1. MDbA module process diagram in BPMN format.

- Sound transmission operating with frequency modulation (FM).

In turn, ancillary retransmission services are also analyzed, namely:

- Analog TV retransmission (RTV);
- Digital TV retransmission (RTVD);
- FM retransmission (RTRFM).

These services are detailed in Tab. II [20].

Several observations about broadcasting services deserve attention. The OM sound broadcasting service operates with three coverage modes: national, regional, and local. They primarily differ from the transmission power utilized by stations, leading to varying degrees of coverage. However, due to the superior service quality offered by FM operations, a migration process to FM sound broadcasting was started in 2013 [21], initially restricted to local OM services. Nevertheless, to ensure parity in broadcasting, non-local operations could also transition to FM operations [21]. Another important transition process refers to the move from analogue to digital TV services (TV to GTVD), which has begun under the provisions of decree 5.820 of 2006 [22].

In addition to their service, channels are categorized based on their intended purpose. Brazilian legislation recognizes four types of broadcasting channel purposes: commercial, educational, public, and public educational [1]. In turn, the Mosaico platform includes geographical data concerning the installation sites of the base stations corresponding to each channel. This data encompasses the municipality's name and code, state (Federative Unit - FU), and geographic coordinates. To ensure the accuracy of this information, the MDbA module utilizes databases supplied by the Brazilian Institute of Geography and Statistics (IBGE). These databases contain shapefiles (`.shp` format) [23] delineating the geographic boundaries of municipalities and FUs across Brazil [24]. Furthermore,

TABLE I
SUMMARY OF BROADCAST CHANNEL STATUS.

Status Code	Status Description
C0	Vacant channel
C1	Granted channel, Awaiting radio frequency act
C2	Granted channel, Awaiting station data
C3	Granted channel, Awaiting licensing
C4	Licensed channel
C5	Fistel generated, Channel pending grant
C7	Expired radio frequency, Awaiting radio frequency renewal act
C8	Licensed channel, Awaiting contract data
C99	Suspended channel
CE-01	Deleted channel

Anatel provides technical details regarding the operation of station equipment, such as antennas and transmission lines. Operational specifics like frequency, effective radiated power (ERP) and radiation pattern utilized in transmissions are also documented in Mosaico.

The channels are linked to a collection of documents pertaining to its procedural cycles, issued by governmental bodies such as the National Congress (NC), Anatel, and MCom itself. These documents serve multiple purposes, including validating transmission rights, renewing or revoking granted licenses, and addressing fee requests and irregularities. The following document attributes are particularly important for the MDbA data processing: type, reason, nature, publication date, and signature date. Access to this channel document data on the Mosaico platform (restricted to permitted users) enables audits and validation of multiple legislative rules applied to broadcasting grants. Additionally, a channel's document collection can be categorized into two sections. The first section, termed as historical documents (HD), constitutes the foundational records of the channel, encompassing all documents issued to it over time. The second section, labeled grant documents (GD), comprises documentation pertaining to grants, including contracts and licenses.

B. Detection of Inconsistencies and Errors in Channel Data

The MDbA module employs the object-oriented programming (OOP) paradigm to manage Mosaico channel records. The approach involves creating distinct classes for various data types, facilitating efficient processing and the implementation of tailored functions. Consequently, individual blocks of information are instantiated with their specific properties. Furthermore, the design encompasses processes dedicated to identifying and rectifying errors within Mosaic records. These discrepancies are logged within a specialized class known as `MosaicoStatus`, which comprises a numerical code and a message that describes the nature of the detected record error.

An important data class to highlight is `MosaicoID`, which serves as a central set for channel-related information and is uniquely identified by the CID. This class consolidates all data associated with a specific channel, bringing together various subclasses that encompass the general data fields within the Mosaico platform. Additionally, it incorporates specific entries and functions, forming a mapping of a channel's information. By grouping this information, it becomes possible to cross-reference multiple datasets to identify inconsistencies. Through the utilization of this data class, errors detected across all data fields of a broadcast channel can be aggregated into instances of the `MosaicoStatus` type. Each instance provides a clear description of the error detected.

There are two categories of rules for data validation implemented in the MDbA module: simple rules (SR) and business rules (BR). The SRs are related to elementary checks regarding the correctness of data fields. On the other hand, the BRs involve legislative and procedural details regarding a channel. These were raised in the DU phase with the help of the specialized MCom team. In Tab. III, the SRs implemented in MDbA are described. For clarification, some observations about the listed SRs can be made. The checks carried out according to the rules between SR-2 and SR-9 apply IBGE databases that include information about municipalities and FUs in Brazilian territory. In turn, validation of an NRLE registration in SR-10 is performed utilizing Brazil's Federal Revenue database. Rules SR-11 and SR-12 apply dictionaries of the types and ratios of documents involved in broadcasting processes. In turn, validations between SR-15 and SR-17 are conducted to verify the dates of signature and publication of channel-related documents.

Table IV provides a summary of the BRs integrated into the MDbA. These rules delineate procedural requirements across various fields of channel data, which involve scenarios where specific criteria must not be violated. If a non-compliance occurs, an error alert is triggered for the respective channel under analysis.

C. Generation of Grant Data

As previously outlined, the GMS requires a new data model tailored to broadcasting license management. This structure must encompass all information necessary to evaluate and manage each broadcasting license. Integrating data from diverse sources is essential to developing an agile and effective GMS.Therefore, grant data generation is initiated with the aggregation phase, in which multiple channels data are joined to form a grant data structure (GDS). As illustrated in Fig. 2, the consumption of information through GDSs speeds up and facilitates grant management due to the consumption of information from a unified source.

According to the knowledge obtained during the DU phase, there are two types of GDSs: composite and simple grants. Composite GDSs are generated by the aggregation of two active channels with distinct services and encompass licenses relating to transition processes, such as migrations and digitalizations. The service combinations allowed in composite GDSs are as follows:

1) TV channel and GTVD channel;
2) OM channel and FM channel originated from a migration process;
3) RTV channel and RTVD channel.

TABLE II
SUMMARY OF THE BROADCASTING SERVICES.

Service Indicator	Service Number	Service Description
TV	248	Broadcasting service intended for the transmission of sounds and images in a television analog channel
FM	230	Sound broadcasting service that operates with frequency modulation (FM) in the band from 76 MHz to 108 MHz
OM	205	Sound broadcasting service that operates with amplitude modulation (AM) in the bands from 525 KHz to 1605 KHz and 1605 KHz to 1705 KHz
GTVD	247	Broadcasting service intended for the transmission of sounds and images in a television digital channel
RTV	800	Service designed to simultaneously retransmit the signals from the analog television generating station
RTRFM	805	Service designed to retransmit in the municipalities of the Legal Amazon (except capitals), simultaneously, the FM station signals from the state capital
RTVD	801	Service designed to simultaneously retransmit the signals from the digital television generating station

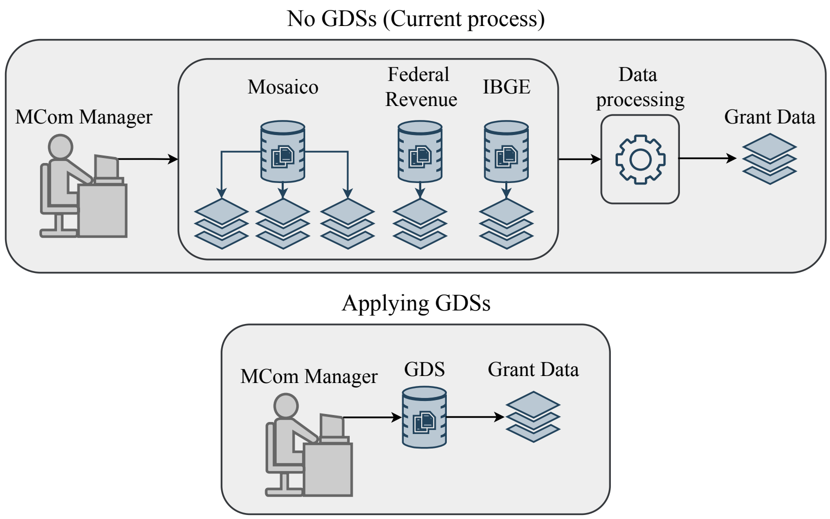


Fig. 2. Illustration of grant management without the use of GDSs and with the application of GDSs.

TABLE III
SUMMARY OF THE SIMPLE RULES IMPLEMENTED IN MDBA.

Rule Code	Rule Description
SR-1	Check if the field is empty
SR-2	Check the validity of geographic coordinates
SR-3	Check validity of FU registration
SR-4	Check validity of municipal registration
SR-5	Check validity of municipal code registration
SR-6	Check the correspondence between the name and the municipality code
SR-7	Check correspondence between the municipality and the FU
SR-8	Check the correspondence between geographic coordinates and FU
SR-9	Check the correspondence between geographic coordinates and city
SR-10	Check the validity of the NRLE registration
SR-11	Check the validity of the Fistel registration
SR-12	Check the validity of the channel purpose registration
SR-13	Check the validity of the status channel code
SR-14	Check the validity of the document type
SR-15	Check the validity of the document reason
SR-16	Check the validity of the document signature date
SR-17	Check the validity of the document publication date
SR-18	Check if the publication date precedes the document signing date
SR-19	Check if the document has invalid dates
SR-20	Checks if the document signature date precedes 01/01/1900
SR-21	Checks if the document publication date precedes 01/01/1900

In turn, simple grants consist of a single active channel related to an arbitrary service. Both simple and composite grants may incorporate data from an indefinite number of inactive channels as long as the services of the inactive channels coincide with the active services of the grant. In addition to service information, the aggregation of channel data in GDSs must incorporate NRLE and municipal code (MC) data. This is because all aggregated data within a given GDS needs to be associated with the same NRLE and MC. This rule ensures the unification of ownership, which is attributed to a specific entity with a designated NRLE; and location information within a GDS. Fig. 3 illustrates the process of aggregating channel data into a composite GDS.

Let us formally define the channel aggregation process. Take $\mathcal{M}$ as the set of all channels contained in the Mosaico and $c \in \mathcal{M}$ is a specific channel instance, identified by a specific CID. Furthermore, assume that the $\mathrm{NR}(c)$ and $\mathrm{MC}(c)$ functions extract the NRLE and MC of a given channel, respectively; and that the indicator function $u(c)$ returns 1 if the channel is active and 0 otherwise. Therefore, the set of channels that share the same NRLE and MC is:

$$\mathcal{C}(s_{\mathrm{NR}}, s_{\mathrm{MC}}) = \{c \in \mathcal{M} : \mathrm{NR}(c) = s_{\mathrm{NR}} \wedge \mathrm{MC}(c) = s_{\mathrm{MC}}\}, \quad (1)$$

where s_{NR} and s_{MC} are arbitrary NRLE and MC sequences and $\wedge$ is the logical conjunction operator. In turn, the subsets

TABLE IV

BUSINESS RULES IMPLEMENTED IN MDBA. THE RULES CHECK FOR PROHIBITED SITUATIONS AND AN ERROR CALL IS RAISED IF THEY OCCUR.

Rule Code	Rule Description
BR-1	Channels with commercial purposes cannot belong to the Federal Union
BR-2	Channels associated with a commercial NRLE cannot have a public purpose
BR-3	An entity, with a given NRLE, cannot have channels with different purposes
BR-4	Retransmission channels with status in {C1,C2,C3,C4,C8} must not have a "grant" reason ordinance in the HDs section
BR-5	Retransmission channels with status in {C1,C2,C3,C4,C8} must have a "grant" reason ordinance in the GDs section
BR-6	TV channels (TV or GTVD) with status in {C1,C2,C3,C4,C8} must have a legislative decree deliberating from the NC in the GDs section
BR-7	TV channels (TV or GTVD) with status in {C1,C2,C3,C4,C8} must not have a legislative decree deliberating from the NC in the HDs section
BR-8	Channels belonging to the Federal Union cannot have C5 status
BR-9	Channels with public purposes must not have status code C5
BR-10	Channels with status in C2 or C3 must have an authorization act for the use of radio frequency in their HDs sections
BR-11	Channels with status code in C5 must have a document with reason "Grant" reason in its HDs section
BR-12	Channels with status code in C5 must not have a document in the GDs section
BR-13	Channels with status in {C1,C2,C3,C4,C8} must have at least one decree in its HDs section
BR-14	TV or GTVD channels with status code in {C1,C2,C3,C4,C8} must have a document of type "legislative decree" and reason "deliberation of the NC" in the GDs section
BR-15	Public purpose channels must have a document of the type "ordinance" and reason "Consignment" in the GDs section
BR-16	Channels with status in C2 must not have a radio frequency use authorization document in their HDs sections
BR-17	RTV channels must not have status in C8
BR-18	Channels with status in C8 cannot have a public purpose
BR-19	Channels with status in C4 must have contract information
BR-20	Channels with status in C3, C4, C7 or C8 must contain information about the transmission ERP
BR-21	Channels with status in C4, C7 or C8 must contain information about the transmission equipment

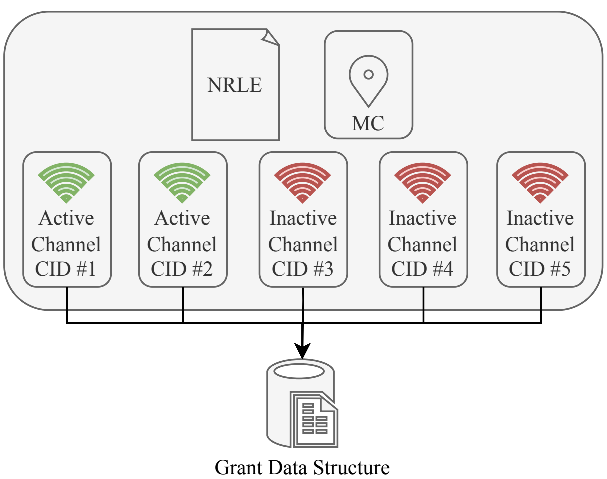


Fig. 3. Illustration of channel data aggregation in a composite GDS, note that all channels share the same NRLE and MC.

of active and inactive channels in $\mathcal{C}(s_{\mathrm{NR}}, s_{\mathrm{MC}})$ are written as

$$\mathcal{C}_1(s_{\mathrm{NR}}, s_{\mathrm{MC}}) = \{c \in \mathcal{C}(s_{\mathrm{NR}}, s_{\mathrm{MC}}) : u(c) = 1\}, \quad (2)$$

and

$$\mathcal{C}_0(s_{\mathrm{NR}}, s_{\mathrm{MC}}) = \{c \in \mathcal{C}(s_{\mathrm{NR}}, s_{\mathrm{MC}}) : u(c) = 0\}, \quad (3)$$

respectively. Finally, the set $\mathcal{S}(s_{\mathrm{NR}}, s_{\mathrm{MC}})$ represents the list of services for the channels contained in $\mathcal{C}(s_{\mathrm{NR}}, s_{\mathrm{MC}})$.

The aggregation proceeds by extracting subsets from $\mathcal{C}(s_{\mathrm{NR}}, s_{\mathrm{MC}})$ that have services corresponding to compound and simple GDSs. As previously explained, composite grants are created from the combination of channels with the following services: $\{\mathrm{TV}, \mathrm{GTVD}\}$, $\{\mathrm{OM}, \mathrm{FM}^*\}$ and $\{\mathrm{RTV}, \mathrm{RTVD}\}$, where FM^* denotes a FM service from a migration channel. In turn, simple grants are generated by a single channel with any service. Therefore, ten grant reference subsets can be defined as $\bar{\mathcal{S}}_m$, $1 \leq m \leq 10$. The first three subsets correspond to service relationships in composite grants, i.e. $\bar{\mathcal{S}}_1 = \{\mathrm{TV}, \mathrm{GTVD}\}$, $\bar{\mathcal{S}}_2 = \{\mathrm{OM}, \mathrm{FM}^*\}$ and $\bar{\mathcal{S}}_3 = \{\mathrm{RTV}, \mathrm{RTVD}\}$. The subsequent subsets, with $m \geq 4$, are of a single element referring to each of the services presented in Tab. II. Fig. 4 presents the pseudocode of the algorithm for aggregating multiple channels in a grant set that represents a GDS.

The procedure explained in Fig. 4 returns a list with N_{G} GDSs denoted as sets $\mathcal{G}_n$, where each set $\mathcal{G}_n$ contains at least one active channel (simple grants). Therefore, the reference date of the n-th grant is calculated as the oldest channel reference date among the elements of $\mathcal{G}_n$. From this reference date, the grant expiration date and its number of renewal cycles can also be calculated. According to Brazilian broadcasting legislation, the period of validity of a grant depends on its service: image and sound transmission services (TV and GTVD) must renew their license every 15 years, while sound services (OM and FM) every 10 years. Ancillary retransmission services (RTR, RTV and RTVD) do not expire, thus, do not have renewal periods. From the reference date and the renewal period, it is possible to calculate the current renewal cycle index of a grant:

$$I_{\mathrm{r}}(\mathcal{G}_n) = \left\lfloor \frac{A(\mathcal{G}_n)}{T_{\mathrm{s}}} \right\rfloor, \quad (4)$$

where $A(\mathcal{G}_n)$ is the grant age (expressed in years), T_{s} is the corresponding renewal period and $\lfloor\cdot\rfloor$ is the floor function. Based on the reference date and the renewal period, it is also possible to calculate the next expiration date of a grant.

It is important to highlight that the aggregation of channel data into a grant structure is an unprecedented process and, therefore, GDSs do not have a predefined GID. To resolve this limitation, in the DU phase, the MCom team suggested the following format for the grant ID:

$$\mathrm{GID} = \{\texttt{yyyymmdd}\}\text{-}\{\texttt{ccccccc}\}\text{-}\{\texttt{sss}\}\text{-}\{\texttt{v}\}. \quad (5)$$

Input: Non-empty set $\mathcal{C}(s_{\mathrm{NR}}, s_{\mathrm{MC}})$
Result: List of grants sets $\{\mathcal{G}_n\}$
Initialize $\mathcal{C}_0(s_{\mathrm{NR}}, s_{\mathrm{MC}})$ and $\mathcal{C}_1(s_{\mathrm{NR}}, s_{\mathrm{MC}})$;
Initialize $\mathcal{S}(s_{\mathrm{NR}}, s_{\mathrm{MC}})$;
Initializes the reference service sets $\bar{\mathcal{S}}_m, m \in [1, 10]$;
Initialize the grants index: $i = 0$;
for $\bar{\mathcal{S}}_r := \bar{\mathcal{S}}_1$ *to* $\bar{\mathcal{S}}_{10}$ **do**
  while $\bar{\mathcal{S}}_r \subset \mathcal{S}(s_{\mathrm{NR}}, s_{\mathrm{MC}})$ **do**
    Initialize a empty grant set: $\mathcal{G}_i = \varnothing$;
    Extract the first $\bar{N}_r = \mathrm{card}(\bar{\mathcal{S}}_r)$ active channels from $\mathcal{C}_1(s_{\mathrm{NR}}, s_{\mathrm{MC}})$ with services given by $\bar{\mathcal{S}}_r$ and allocate in $\mathcal{K}_1$;
    Extract all inactive channels from $\mathcal{C}_0(s_{\mathrm{NR}}, s_{\mathrm{MC}})$ with services given by $\bar{\mathcal{S}}_r$ and allocate in $\mathcal{K}_0$;
    Add the channels to the grant set: $\mathcal{G}_i \leftarrow \mathcal{G}_i \cup \{\mathcal{K}_0 \cup \mathcal{K}_1\}$;
    Remove the channels from $\mathcal{C}(s_{\mathrm{NR}}, s_{\mathrm{MC}})$: $\mathcal{C}(s_{\mathrm{NR}}, s_{\mathrm{MC}}) \leftarrow \mathcal{C}(s_{\mathrm{NR}}, s_{\mathrm{MC}}) \setminus (\mathcal{K}_0 \cup \mathcal{K}_1)$;
    Recalculate the set of services $\mathcal{S}(s_{\mathrm{NR}}, s_{\mathrm{MC}})$ based on the updated $\mathcal{C}(s_{\mathrm{NR}}, s_{\mathrm{MC}})$;
    Update the grants index: $i \leftarrow i + 1$;
  end
end
Get the number of processed grants: $N_{\mathrm{G}} = i$;
return $\{\mathcal{G}_n\}, 0 \leq n \leq N_{\mathrm{G}} - 1$;

Fig. 4. Channel data aggregation algorithm in GDSs.

The sequence in (5) is a string with 17 numerical digits. The first part refers to the reference date of the grant, with eight digits, followed by the seven digits sequence of the municipality code and the three digits service code (*c.f.* Tab. II). The last digit (v) is a check digit, calculated according to Lunh's algorithm [25] and used to attest to the correctness of the GID.

D. Limits of Decree 236 and Law 14.812

Decree 236 of 1967 limits the ownership of broadcasting licenses by an entity and its partners [4]. The original text delimits the number of grants per service, including limitations on coverage in OM broadcasting. However, the text was modified in Law 14.812 of January 2024, revoking the coverage limitations [26]. In this new version, limits are imposed only by the type of media. Therefore, entities holding licenses, and their partners, must have a maximum of 20 sound and image broadcasting licenses (TV and GTVD) and 20 sound broadcasting licenses (OM and FM). Ancillary retransmission (RTV, RTVD and RTRFM) services do not count towards the limits of the decree. It is important to highlight that limits are imposed not only on legal entities, but also on their partners.

From the GDSs generated in MDbA, it is possible to automatically check whether the entities and their partners respect the limits of decree 236. Using Brazil's Federal Revenue database, MDbA extracts the corporate details of entities, which may contain individuals or legal entities, and associates them with a grant. The extracted corporate tables have multiple levels and take into account the partners of other partner entities in a recursive manner, resulting in a corporate tree linked to a grant, as illustrated in Fig. 5. Therefore, all elements of a corporate tree are seen as participants in the grant corresponding to the root entity. As a result, the count carried out by MDbA takes into account all members of Brazilian broadcasting's corporate trees.

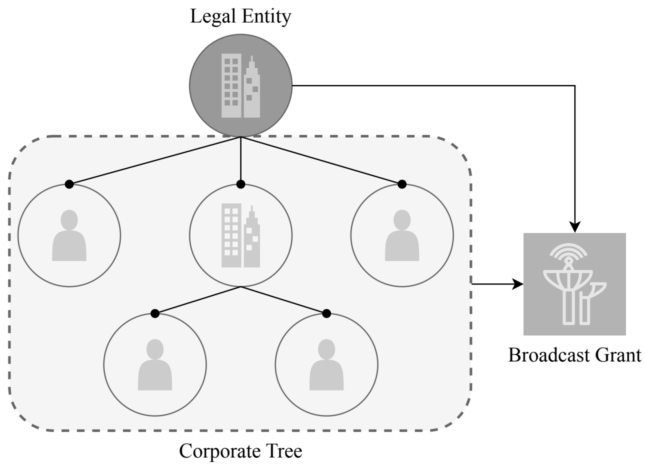


Fig. 5. Illustration of the relationship between an entity's corporate tree and a grant.

III. Numerical Results

This section presents numerical results relating to error detection, generation of grant data and validation of the limits of decree 236. During the error detection and inconsistency identification phase of MDbA processing, the rules outlined in Tab. III and Tab. IV were applied to the entire Mosaico database, which consists of 46941 channels, comprising 27678 active and 19263 inactive channels[2]. It is important to emphasize that a large scale analysis like this would be impossible without the automated process that this work develops. Thus, the following results are real examples of how the MCom management can considerably improve the quality of their duty, i.e. broadcasting grants supervision.

This study focuses on presenting detection results exclusively for active channels, representing 58.9% of the Mosaico database. Utilizing MDbA processing, a total of 46941 registration errors were identified among the 27678 active channels, averaging 2.29 errors per channel, with a standard deviation of 2.38. The maximum number of errors observed in one active channel is 74. The total numbers of errors based on SRs and BRs are 23376 and 23565, respectively. The rule that resulted in the highest number of errors was SR-17, with 13909 occurrences. SR-17 specifically verifies the accuracy of publication dates in various channel documents. If these dates are incorrect or missing, an error is flagged. Additionally, BR-5 was the most error-prone BR, with 7188 instances. This rule verifies whether a retransmission channel with a specific status (*cf.* Tab. IV) contains an ordinance with the reason "grants" in its GDs section. If such an occurrence is detected, the system generates the corresponding error code. A summary of the error detection results is presented in Tab. V.

[2] Values considering a dump of the Mosaico database dated 03/07/2024.

TABLE V
GENERAL SUMMARY OF REGISTRATION ERROR DETECTION IN THE MOSAICO DATABASE.

Description	Value
Total number of channels	46941
Number of active channels	27678
Number of errors on active channels	46941
Average number of errors per active channel	2.29
Max. number of errors in one channel	74
Number of errors based on SRs	23376
Number of errors based on BRs	23565
Simple rule that raised the most errors	SR-17 (13909 cases)
Business rule that raised the most errors	BR-5 (7188 cases)

Fig. 6 shows the frequency of occurrence of errors related to SRs. It's worth noting that SRs without value bars indicate error-free records, signifying their validity. Based on SR-1, 2190 empty fields were found in the active channel data. The fields that require the most completion refer to the type and reason for the channel documents. In other hand, it can be observed that the records of geographic coordinates, state (FU) and municipality of the channels, verified in rules SR-2, SR-3, SR-4 and SR-5, were individually recorded correctly. Moreover, the correspondence between FU and municipality, examined by SR-7, proves accurate across all active channels. However, certain geographical correlations display errors, such as those highlighted by the SR-6 rule, revealing four channels with inconsistent municipal name and code records. Furthermore, the correspondence between geographic coordinates and state, examined in rule SR-8, exposes 95 errors. In these cases, the positions indicated by the coordinates fall outside the geographic limits of the registered state.

Fig. 7 presents a heat map illustrating the distribution of errors detected by the SR-8 rule across Brazilian states. Notably, the state of Minas Gerais stands out with the highest number of non-corresponding coordinates, totaling 17 cases. For better visualization, Fig. 8 displays the inconsistent coordinates in relation to the state of Minas Gerais. The presented points indicate discrepancies, ranging from minor geographical errors near state borders to some points located considerably far from the border polygon, including geographic points in the ocean. In turn, the validation of the correlation between geographic coordinates and municipalities carried out in the SR-9 rule raised 889 errors, that is, 889 channels have coordinate records outside the limits of the registered municipality.

Continuing the analysis of the SRs, according to the results of the SR-10 and SR-12 rules presented in Fig. 6, it can be observed that the NRLE and channel purpose records do not contain any errors. Regarding the Fistel code, the SR-11 implementation identified 11 empty records pertaining to channels with diverse statuses in `C1`, `C4`, `C5`, and `C7`. Furthermore, SR-13 detected only two instances of channel status records deviating from the prescribed formatting. In these cases, the status was incorrectly recorded as FM-`C8`, incorporating an unnecessary service prefix. The correct format, standardized in MDbA, is `C8` without any additional prefix. Conversely, SR-14 processing did not reveal any inconsistencies in the records of channel document types. However, SR-15 identified 893 errors in the reasons provided for the documents. These errors are

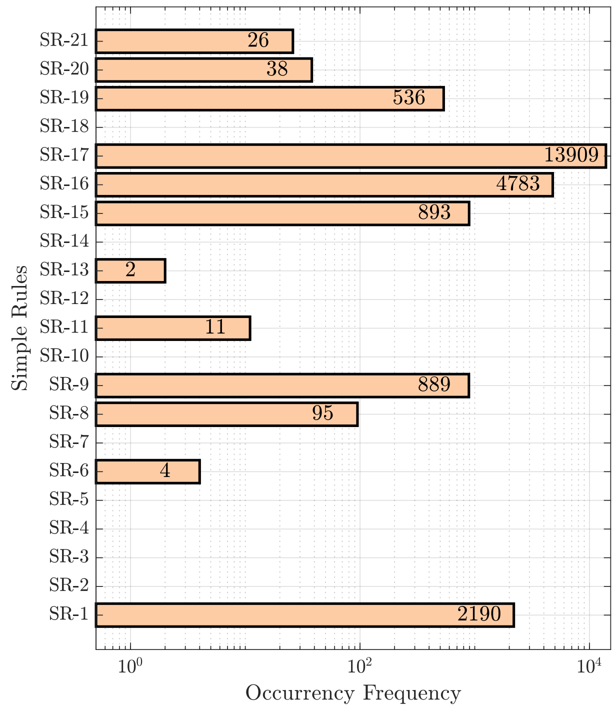


Fig. 6. Frequency of occurrence of implemented SRs. SRs without value bars did not raise any errors.

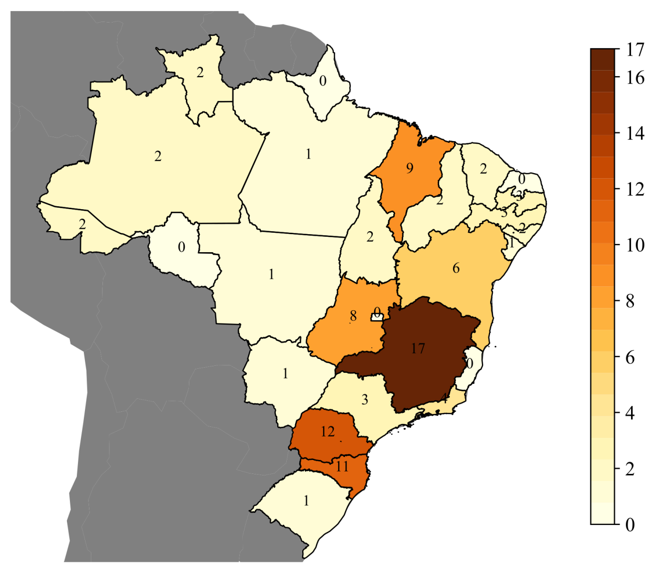


Fig. 7. Heat map along the Brazilian states of the number of inconsistencies in coordinate records based on the SR-8 rule.

characterized by deviations from a predefined dictionary of reference reasons.

The most prevalent errors identified by SRs across the channels pertain to document dates, as assessed by rules SR-16 through SR-21. Specifically, SR-16 and SR-17 validate the signature and publication dates of documents, respectively, resulting in a total of 18,692 cases of errors detected by these rules. In turn, the SR-18 processing did not find any key

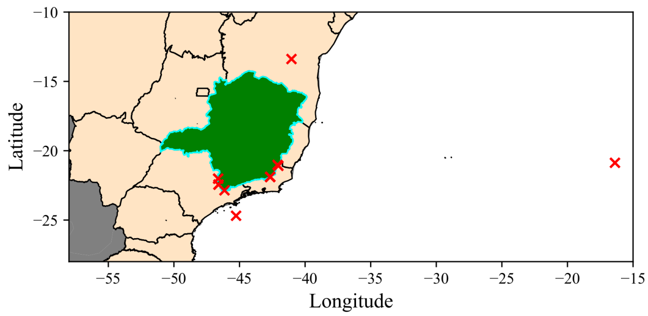


Fig. 8. Illustration of inconsistent coordinates for the Brazilian state of Minas Gerais found by implementing rule SR-8.

documents dates that precede the channel reference date[3]. Furthermore, the SR-19 rule, designed to identify documents with both incorrect signature and publication dates by intersecting SR-16 and SR-17, detected 536 errors. Finally, validations performed by SR-20 and SR-21 revealed 38 signature dates and 26 publication dates preceding the year 1900, displaying likely typos in your records. It's noteworthy that the high number of errors in document dates, amounting to 19,292 cases, can be attributed to the substantial volume of documents per channel. For instance, the most densely populated channel contains 94 documents.

Fig. 9 provides an overview of the error detection results based on BRs. As previously elucidated, BRs are rules tailored to the specificities of the channels' life cycle, allowing for the validation of channel procedures. According to the findings corresponding to BR-1, no active channel with commercial purposes is associated with the Federal Union. Conversely, BR-2 processing revealed that five channels owned by commercial legal entities serve public purposes, contravening Brazilian channel management standards. Additionally, BR-3 processing detected 745 cases of entities operating channels with divergent purposes. Validation of rules BR-4 and BR-5, which pertain to retransmission channels with specific statuses containing documents in incorrect sections, yielded a total of 7265 detections. Notably, BR-5 emerges as the most frequently violated business rule. Furthermore, validation of rules BR-6 and BR-7, applied to sound and image transmission services (TV or GTVD), identified a combined total of 901 errors.

Validations conducted based on BR-8 and BR-9 rules reveal that a total of 366 public channels, either by virtue of their purpose or their association with the Federal Union, possess `C5` status, a condition prohibited by channel legislation. Moreover, the validation between rules BR-10 and BR-16, which address the presence of specific documents in the HDs and GDs sections of particular channels, identified a total of 7189 inconsistencies. Notably, the BR-16 assessment was the second most error-prone, exposing 5994 channels with inconsistencies. Checks pertaining to BR-18 and BR-19 did not yield any errors, signifying that all channels in status `C8` lack public purpose and all channels in `C4` provide contract information, respectively. Lastly, verification of technical information from

[3]The SR-18 processing does not take into account all documents present in the HD section, taking into account only a set of documents with specific types and reasons.

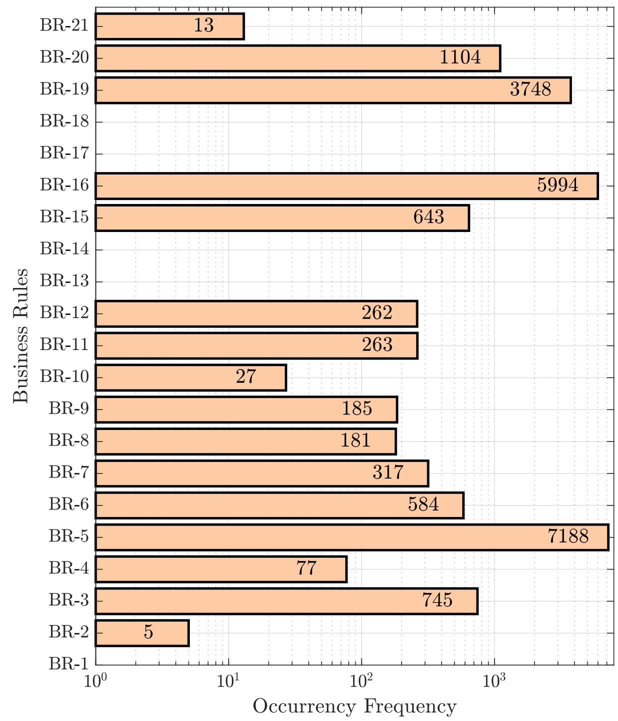


Fig. 9. Frequency of occurrence of implemented BRs. BRs without value bars did not raise any errors.

transmitters on channels with specific statuses, as per BR-20 and BR-21 rules, revealed 1117 errors.

Focusing now on the results relating to the generation of grants sets, the MDbA processing returned a total of 23257 GDSs, comprising 15484 (67.05%) simple grants and 7773 (32.95%) compound grants. Among these, 5630 (22.5%) grants pertain to transmission services, while the remaining 17627 (77.5%) GDSs relate to ancillary retransmission services. The number of grants per service is detailed in Tab. VI. Retransmission grants are the most prevalent, with RTVD constituting 36.5% of Brazilian grants, and compound grants of type RTV plus RTVD making up 30% of Brazilian broadcasting. Notably, transmission services primarily comprise FM and OM, accounting for 17.4% and 2.47% respectively. Fig. 10 depicts the distribution of grants across Brazilian territory. The visualization highlights a notable concentration of broadcasting licenses in the southeastern states, particularly in Minas Gerais, accounting for 3928 licenses. In contrast, Amapá exhibits the lowest concentration of grants, with only 108 licenses.

Fig. 11 shows the histogram of broadcasting grants reference years, computed using the reference dates established by the MDbA. A notable accumulation of licenses is evident from the 1980s onwards. Particularly, the period with the highest number of granted licenses spans from 2020 to 2021, totaling 2669 grants. Additionally, Fig. 12 exhibits the empirical cumulative distribution function (CDF) of the grants expiration years. This analysis reveals that the majority of grants expire by 2028. Moreover, more than 90% of currently active grants are projected to expire by 2033. Fig. 13 presents a histogram of the renewal cycle index for broadcasting grants.

TABLE VI
NUMBER OF GRANTS PER SERVICE.

Grant type	Service	Value
Compound Grants	TV + GTVD	452
	FM + OM	348
	RTV + RTVD	6973
	Total	7773
Simple Grants	TV	26
	FM	4036
	OM	574
	GTVD	194
	RTV	2051
	RTRFM	106
	RTVD	8497
	Total	15484

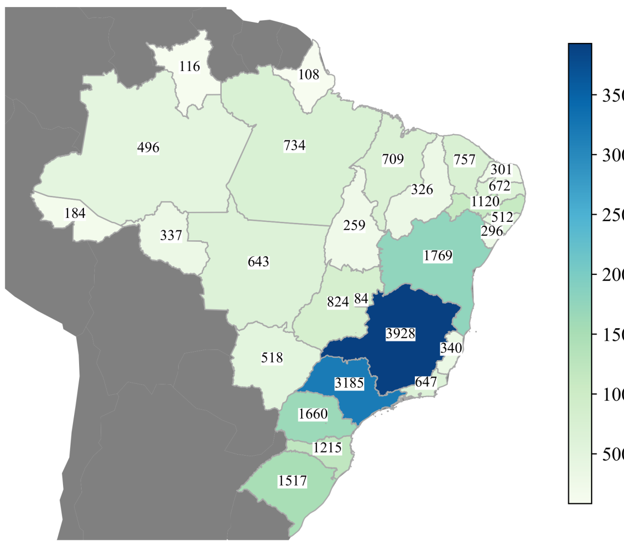


Fig. 10. Heat map of the number of grants along the Brazilian states.

Initially, it is noticeable that 9.7% of broadcasting licenses remain in their first cycle (index 0) and have never undergone a renewal process. Additionally, the mode of the renewal cycle index is concentrated at index 1, in which 36.1% of grants having undergone renewal once. Finally, it is evident that grants exceeding the fourth renewal cycle comprise less than 1% of the total.

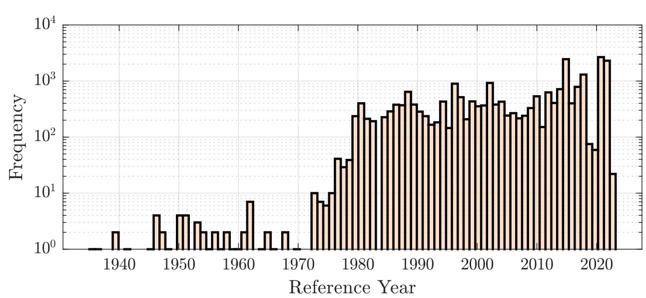


Fig. 11. Histograms of the grants reference years.

Finally, to evaluate the limits of decree 236, considering its modifications to law 14.812, the number of grants from all partners in Brazilian broadcasting corporate trees was counted. In Fig. 14, a histogram for the distribution of grants per partner is presented, emphasizing the upper limit established in the revision by the Law 14.812. This limit restricts ownership to a maximum of 20 licenses for sound and image transmission (TV and GTVD) and 20 licenses for sound broadcasting (FM and OM), totaling a maximum of 40 grants. According to the results of the MDbA processing, no partner violates the limits of decree 236, in which the partner closest to the limit has 17 grants.

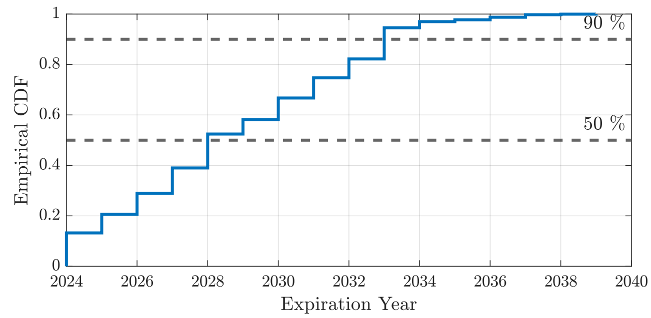


Fig. 12. Empirical CDF of the grants expiration years.

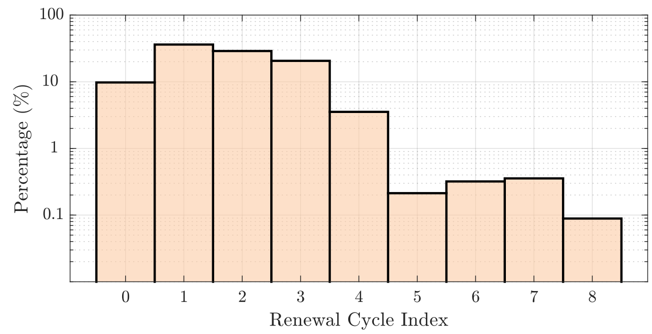


Fig. 13. Histogram of the renewal cycle index.

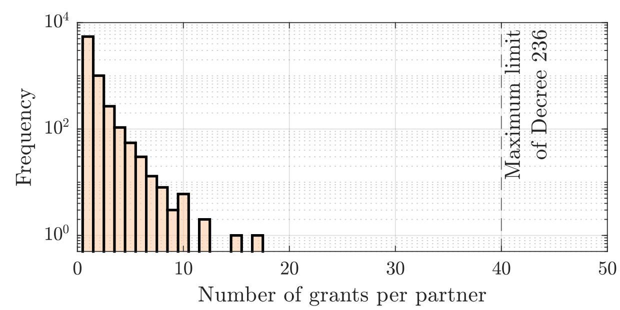


Fig. 14. Histogram of the number of grants per partner. The upper limit determined by decree 236 of 40 grants (20 TV services + 20 sound services) is highlighted.

IV. CONCLUSION

This work presented the data processing results achieved by the Mosaico database analysis (MDbA) module, which was developed to detect errors in the Mosaico platform for managing Brazilian broadcasting grants. This automated assessment of errors and violations allowed the Ministry of Communications (MCom) to implement appropriate measures for data sanitization and regulation. Besides detecting inconsistencies, the MDbA module also aggregated channel data into a new, efficient structure for grant management. This structure facilitated the extraction of critical information regarding broadcasting grants, such as reference and validity dates. Additionally, a fundamental verification step carried out by the MDbA involved checking for potential violations of the Decree 236.

MDbA was developed specifically to handle data related to broadcasting licenses. Thus, it is not applicable to other business models, such as cellular or satellite communications applications due to the different rules and legal frameworks involved. Moreover, MDbA consumes in a static fashion information from different databases, such as the Federal Revenue Service and the Brazilian Institute of Geography and Statistics (IBGE). Future work involves developing this functionality to consume updated data directly from the source. Based on the work conducted, the future grants management system, still under development, is designed to utilize sanitized data, thereby improving the efficiency and performance of Brazilian broadcasting management. Furthermore, regulation can be adjusted and automated through grant monitoring methodologies, reducing the procedural analysis burden on the MCom.

ACKNOWLEDGMENT

The authors acknowledge MCom team for the partnership during the development of this work, providing key information and data for the achievement of the results.

REFERENCES

[1] C. A. Lopes, "Regulação das Outorgas de Radiodifusão no Brasil – Uma breve análise," Câmara dos Deputados, Tech. Rep., Dec. 2009. [Online]. Available: https://bd.camara.leg.br/bd/bitstream/handle/bdcamara/5404/regulacao_outorgas_lopes.pdf

[2] A. Simis, "A Legislação Sobre as Concessões na Radiodifusão," *UNIrevista*, vol. 1, no. 3, pp. 1–16, Jul. 2006.

[3] M. S. L. F. Lima and M. de Matos Ramos, "Sobre o Uso Eficiente do Espectro Radioelétrico," Ministério da Fazenda, Secretaria de Acompanhamento Econômico (SEAE), Brasília, Tech. Rep., Dec. 2006.

[4] H. de Alencar Castelo Branco, "Decreto-Lei Nº 236, De 28 de Fevereiro de 1967," Online, Presidência da República, Feb. 1967. [Online]. Available: https://www.planalto.gov.br/ccivil_03/decreto-lei/Del0236.htm

[5] ——, "Decreto-Lei Nº 200, De 25 de Fevereiro de 1967," Online, Presidência da República, Feb. 1967. [Online]. Available: https://www.planalto.gov.br/ccivil_03/decreto-lei/del0200.htm

[6] F. H. Cardoso, "Lei Nº 9.472, De 16 de Julho de 1997." Online, Presidência da República, Jul. 1997. [Online]. Available: https://www.planalto.gov.br/ccivil_03/leis/l9472.htm

[7] Mosaico – Sistema Integrado de Gestão e Controle do Espectro. [Online]. Available: http://sistemas.anatel.gov.br/se/public/view/b/srd.php

[8] Federal Communications Commission – FCC, Online, Feb. 2024. [Online]. Available: https://www.fcc.gov/

[9] ——, "Licesing and Management System," Online, Feb. 2024. [Online]. Available: https://enterpriseefiling.fcc.gov/dataentry/public/tv/publicSearchLanding.html

[10] Canadian Radio-Television and Telecommunications Commission – CRTC, Online, Feb. 2024. [Online]. Available: https://crtc.gc.ca/eng/home-accueil.htm

[11] C. Lim, K.-J. Kim, and P. P. Maglio, "Smart Cities with Big Data: Reference Models, Challenges, and Considerations," *Cities*, vol. 82, pp. 86–99, 2018, DOI: 10.1016/j.cities.2018.04.011. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0264275117308545

[12] J. R. Reis, J. Viterbo, and F. Bernardini, "A Rationale for Data Governance as an Approach to Tackle Recurrent Drawbacks in Open Data Portals," in *Proceedings of the 19th Annual International Conference on Digital Government Research: Governance in the Data Age*. New York, NY, USA: Association for Computing Machinery, 2018, DOI: 10.1145/3209281.3209354.

[13] J. S. Saltz, "CRISP-DM for Data Science: Strengths, Weaknesses and Potential Next Steps," in *2021 IEEE International Conference on Big Data (Big Data)*, 2021, pp. 2337–2344, DOI: 10.1109/BigData52589.2021.9671634.

[14] F. Martínez-Plumed, L. Contreras-Ochando, C. Ferri, J. Hernández-Orallo, M. Kull, N. Lachiche, M. J. Ramírez-Quintana, and P. Flach, "CRISP-DM Twenty Years Later: From Data Mining Processes to Data Science Trajectories," *IEEE Transactions on Knowledge and Data Engineering*, vol. 33, no. 8, pp. 3048–3061, 2021, DOI: 10.1109/TKDE.2019.2962680.

[15] G. V. Pereira, G. Eibl, C. Stylianou, G. Martínez, H. Neophytou, and P. Parycek, "The Role of Smart Technologies to Support Citizen Engagement and Decision Making," *IGI Global*, vol. 20, no. 4, pp. 1–17, 2018, DOI: 10.4018/978-1-6684-3706-3.ch036.

[16] R. Wirth and J. Hipp, "CRISP-DM: Towards a Standard Process Model for Data Mining," in *Proceedings of the 4th International Conference on the Practical Applications of Knowledge Discovery and Data Mining*, Manchester, UK, 2000, pp. 29–39.

[17] Object Management Group Business Process Model and Notation, Online, Nov. 2024. [Online]. Available: https://www.bpmn.org/

[18] H. de Alencar Castelo Branco, "Lei Nº 5.070, De 7 de Julho de 1966." Online, Presidência da República, Jul. 1997. [Online]. Available: http://www.planalto.gov.br/ccivil_03/leis/l5070.htm

[19] V. G. P. C. Geeverghese, "O Impacto da Qualidade Regulatória da Anatel no Setor de Telecomunicações," Master's thesis, Escola Brasileira de Administração Pública e de Empresas, Brasília, Brazil, June 2022, Available at https://repositorio.fgv.br/server/api/core/bitstreams/9996fe1c-c5b8-43ee-9f61-a48b5afc48c7/content.

[20] Agência Nacional de Telecomunicações (Anatel), "Serviços de Radiodifusão," Online, Jul. 2023. [Online]. Available: https://www.gov.br/anatel/pt-br/regulado/radiodifusao/servicos-de-radiodifusao

[21] R. M. Muniz, G. P. Silva, C. A. M. Gold, D. M. de Oliveira, P. B. de Ávila, and G. L. de Albuquerque Neto, "Proposta de Extinção do Serviço de Radiodifusão Sonora em Ondas Médias com Caráter Local e Adaptação para o Serviço de Radiodifusão Sonora em Frequência Modulada," Secretaria de Serviços de Comunicação Eletrônica – Ministério das Comunicações, Tech. Rep., Jul. 2013.

[22] L. I. L. da Silva, "Decreto-Lei Nº 5.820, De 29 de Junho de 2006," Online, Presidência da República, Jun. 2006. [Online]. Available: https://www.planalto.gov.br/ccivil_03/_Ato2004-2006/2006/Decreto/D5820.htm

[23] Environmental Systems Research Institute (ESRI), "ESRI Shapefile Technical Description," Tech. Rep., Jul. 1998.

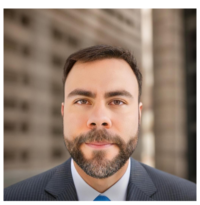
[24] Instituto Brasileiro de Geografia e Estatística (IBGE), "Malhas territoriais." [Online]. Available: https://www.ibge.gov.br/geociencias/organizacao-do-territorio/malhas-territoriais.html

[25] H. P. Luhn, "Computer for verifying numbers," U.S. Patent 2 950 048A, August 23, 1960.

[26] L. I. L. da Silva, "Lei N° 14.812, De 15 de Janeiro de 2024," Online, Presidência da República, Jan. 2024. [Online]. Available: https://www.planalto.gov.br/ccivil_03/_Ato2023-2026/2024/Lei/L14812.htm

Higo Thaian Pereira da Silva was born in Nova Floresta, Paraíba, Brazil, in 1993. He received the B.Sc. degree in electrical engineering in 2016 at the Federal University of Paraíba. At the Federal University of Campina Grande, he received his M.Sc. and D.Sc. degrees in electrical engineering in 2018 and 2023, respectively. His research interests include signal processing and the channel modeling of wireless Communication.

Pedro Henrique Dornelas Almeida received his B.Sc. degree in Communication Network Engineering from University of Brasília, Brazil, in 2023. He received his M.Sc. degrees in electrical engineering in 2024, and he is currently D.Sc student at the same university. His main research interests include digital signal processing and mobile networks.



José Alberto Sousa torres was born in Salvador, Bahia, Brazil, in 1982. He received his Bachelor's degree in IT Systems Analysis in 2005 from the State University of Bahia. At the University of Salvador, he obtained his M.Sc. in Systems and Computing in 2012. He is currently a doctoral student in Electrical Engineering at the University of Brasília. His research interests include signal processing, data science, and theoretical and practical applications of machine learning.

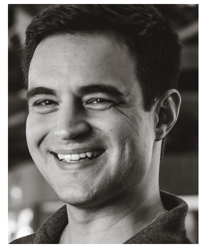
Uanderson Aguiar da Ponte Frota was born in Sobradinho, Brasília, Brazil, in 1982. He graduated in Data Processing in 2007 from the Institute of Exact Sciences (ICEX). In 2011, he obtained a specialization in Software Engineering, followed by a specialization in Project Management in 2015. His research interests include agile methodologies, software design, implementation and evolution, as well as effective methodologies for the development of digital products.

Pedro Duarte Alvim was born in Brasília, DF, Brazil, in 2000. He received the B. Sc. degree in Communication Network Engineering in 2023 at the University of Brasília. He is studying for his M. Sc. degree in Eletrical Engineering with his research interests including channel modeling of wireless communication systems and machine learning.

Daniel Pereira Gonçalves was born in Brasília, Brazil, in 1997. He received the B.Sc. degree in network engineering in 2020 at the Federal University of Brasília. His research interests include Big Data solutions and machine learning applied to wireless communication.

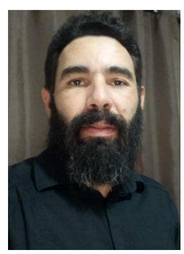
Hugerles S. Silva received his B.Sc., M.Sc. and Ph.D. degrees in Electrical Engineering from University Federal of Campina Grande, Brazil, in 2014, 2016 and 2019, respectively. He is currently an adjunct professor at the University of Brasília. His main research interests include wireless communication, digital signal processing and wireless channel modeling.

Ugo Silva Dias was born in Bel’em, Par’a, Brazil, in 1981. He received the B.Sc. degree in Electrical Engineering from The Federal University of Par’a, Brazil, in 2004, and the M.Sc. and Ph.D. degrees in Electrical Engineering, from The State University of Campinas, Brazil, in 2006 and 2010, respectively. Since March 2010, Dr. Ugo Dias is Professor at University of Brasilia (UnB), Brazil. He is a faculty member of the Department of Electrical Engineering. Prof. Dias has been involved on the Organizing Committee of several conferences. He acts as a Scientific Consultant of the National Council of Scientific and Technological Development (CNPq), Brazil, and he is a Productivity Research Fellow of CNPq. Currently, Prof. Dias serves as President of the Brazilian Telecommunications Society (SBrT) and advisor of the IEEE ComSoc UnB Student Branch Chapter. He is a Senior Member of both IEEE and SBrT, and Member of IEEE Communications Society and Brazilian Communications Committee.



Daniel G. Silva was born in Botucatu, Brazil, in 1983. He received the B.S. degree in computer engineering and the M.S. and Ph.D. degrees in electrical engineering, all from the University of Campinas (UNICAMP), São Paulo, Brazil, in 2006, 2009, and 2013, respectively. Currently, he is a Professor at the Department of Electrical Engineering (ENE) of the University of Brasília (UnB). His main research interests are information theoretic learning and computational intelligence.

Thiago Aguiar Soares has a master’s degree in Electrical Engineering – Telecommunications from the University of Brasília (UnB), a postgraduate degree in Telecommunications Regulation from the National Institute of Telecommunications (Inatel) and a degree in Electrical Engineering from UnB. He worked at the National Telecommunications Agency (Anatel) for 12 years, coordinating projects in digital television, digital radio, spectrum management, technical regulation of broadcasting services, implementation of IT systems, among others. Since 2020, he has been General Coordinator of Innovation, Regulation and Systems at the Secretariat of Electronic Social Communication of the Ministry of Communications, being responsible for developing public policies for the broadcasting sector, including technological and regulatory innovation strategies. Thiago also represents the Brazilian government in international forums and is currently President of Study Commission 6 of the Radiocommunications Sector of the International Telecommunications Union (ITU-R), in which issues related to the global harmonization and standardization of technologies for the radiocommunications sector are discussed.



Thiago Rizza Silva was born in 1982 in Brasília, Federal District, Brazil. He graduated as a Telecommunications Technician from the Technical School of Brasília in 2005. With a 16-year career at the Ministry of Communications (MCom), he is a specialist in broadcasting data, systems and documentation.