CASE STUDY:

Data Transformation
Providing Metadata Assurance to ITV

MetaBroadcast itv

Background

ITV is **an integrated producer broadcaster consisting of ITV Studios and Media & Entertainment** and is the largest commercial broadcaster in the UK. Its production division, ITV Studios, is a leading global creator, producer and distributor of high-quality scripted and unscripted content. With a content catalogue of more than 90,000 hours of active content, ITV Studios operates in 13 countries with over 60 labels, producing content for international television broadcasters and global streaming platforms. ITV's Media & Entertainment division distributes content across its ad-funded linear channels and its ITVX streaming platform. ITV's purpose is to entertain and connect with millions of people in the UK and globally.

Metadata Essential to Strategic Initiatives

ITV's priorities include growing its scripted business, maximising the monetisation of high-value content formats, attracting more monthly active users to ITVX, and maintaining its leadership position in the UK broadcast market. Strategic initiatives to achieve these priorities require using descriptive and collaborative metadata. However, ITV faced the challenge of inconsistent metadata formats spread across siloed repositories, each managed by different lines of business.
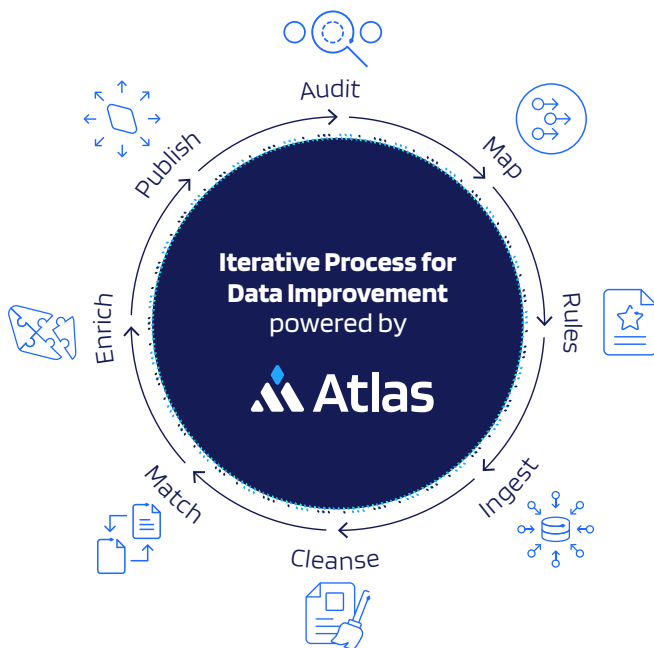
Familiarity with ITV's broadcast-related metadata due to its long-term relationship with UK JIC, Barb, gave MetaBroadcast insight into the issues ITV was facing. MetaBroadcast proposed a two-phase metadata transformation project to consolidate relevant data into a single metadata repository that would reflect:

- Accurate data, including correct titles, spelling, genre unification, runtime, and use of alpha, numeric, or special characters

- Validated data IDs, brands, hierarchies, and genres

- Relevant data that provides context about the content and that is useful to ITV's lines of business

- Removal of irrelevant text used to highlight data status in historic systems

An Iterative Approach to Ensuring Data Quality

MetaBroadcast deployed an iterative, systematic approach to ingest, cleanse, and unify ITV's metadata. The first step was to audit the current status of existing ITV metadata. ITV identified internal data sources and prioritised relevant data fields. Metadata was ingested from three internal ITV platforms into Atlas. Upon ingest, data was deduplicated and mapped against our industry-specific best practices data schema derived from over a decade of unifying, cleansing, and normalising metadata.

Iterative Process for Data Improvement powered by Atlas

Audit · Map · Rules · Ingest · Cleanse · Match · Enrich · Publish

Leveraging automated rules-based processes to identify missing, malformed, misspelt, or incomplete metadata, our audit quickly identified:

- Missing data fields such as genres, titles, brands, images, etc.

- Inconsistencies in metadata formats

- Exceptions such as invalid character types, data sequences, abbreviations

- Lack of consistency of data formats across similar fields

Upon completion of the initial data audit, MetaBroadcast provided ITV with a detailed report quantifying the level of data completeness, accuracy, and consistency and providing a heat map visualising the quality of all data fields. The report also reflected automated enrichment enabled by MetaBroadcast. The audit findings provided clear indicators of where data quality improvements could be made. →
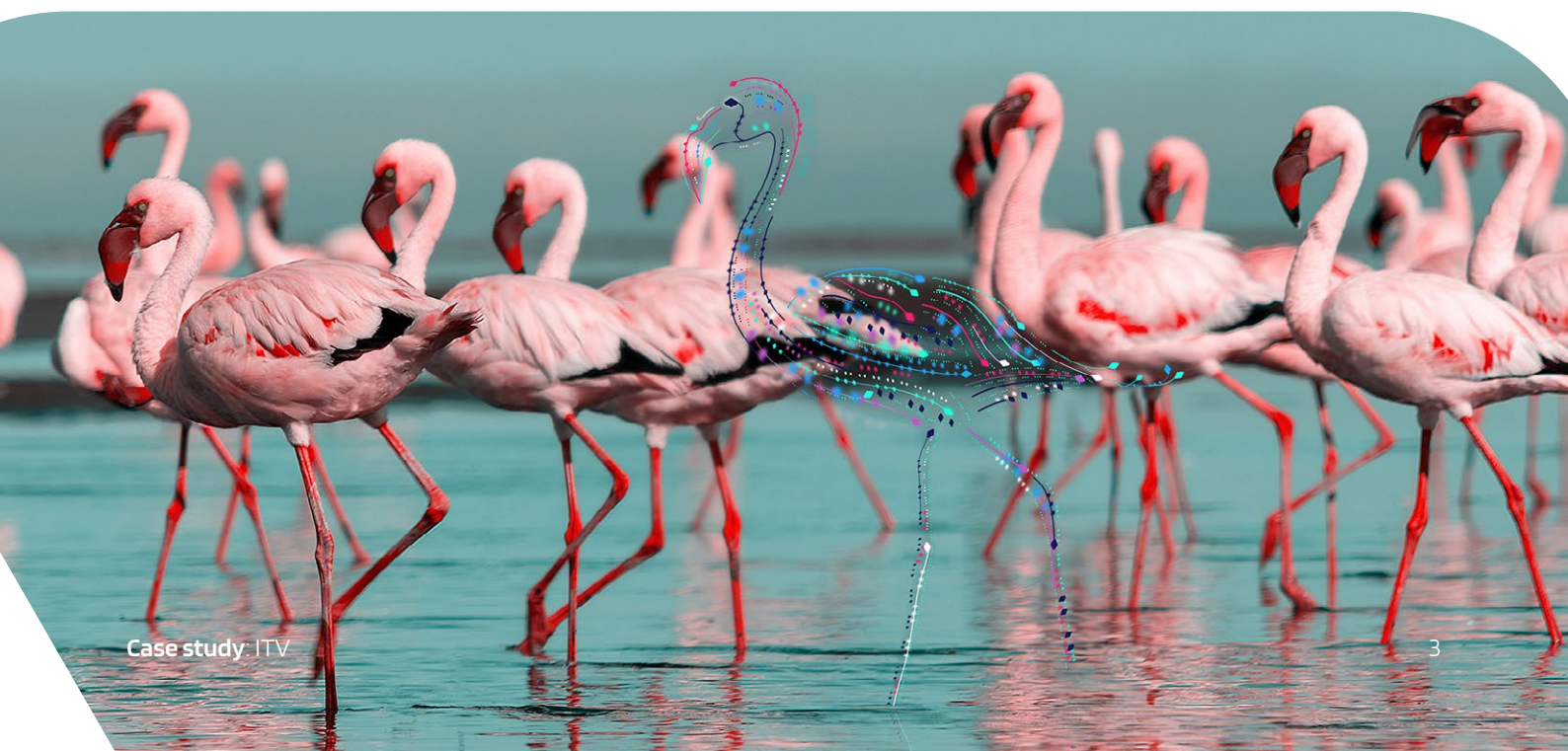
The second step of the engagement was to cleanse the consolidated metadata. Together, MetaBroadcast and ITV agreed on which automated cleansing rules would be applied to ingested metadata. In addition, bespoke rules were defined to provide enhanced cleansing based on the data source.

The resulting clean data is a prerequisite for initiating an automated equivalence process, where content IDs and related content records from ITV's sources are matched. This step leverages MetaBroadcast's expertise in having performed equivalence for our customers for more than a decade. Our metadata repository of over 140 million master MBIDs and their associated content records reflects millions of data fields ingested, cleansed, and equived from many sources (e.g., ITV, Gracenote, IMDb, Press Association, Wikipedia, broadcaster CMS, etc.). The records have been persistently updated as existing data changed or new data became available, giving MetaBroadcast enhanced capabilities to match content records successfully and deliver a consolidated metadata repository to ITV.

Additionally, MetaBroadcast collaborated with ITV to define field-level IDs aligned to each source, allowing for precise referencing and linking of metadata across different systems and processes. This level of granularity facilitates efficient data exchange between platforms while giving ITV a standardised way to access and reference specific pieces of metadata.

A Collaborative Cleansing Process

Throughout the engagement, MetaBroadcast maintained consistent communication with ITV. Weekly and ad-hoc meetings ensured all stakeholders were aware of progress and outcomes. MetaBroadcast provided ITV with reports validating ingest and cleansing rules, audit outcomes, and data mapping targets. MetaBroadcast's expertise in metadata management has resulted in mature, tested data cleansing rules that guide automation.

To address ITV's concerns about the level of automation, hard and soft rules were defined, allowing for the identification of data exceptions that required manual review. The exception review process led to ongoing refinement of the rules. Overall, this iterative process created 40 data cleansing rules to support clear and coherent reports, identifying opportunities to improve data quality and depth.

Key Outcomes

MetaBroadcast helped ITV lay the groundwork for reducing its metadata costs. ITV had acknowledged that a manual data clean-up project would not be cost-effective or timely.

Upon completing the data audit and cleansing, MetaBroadcast quickly and easily mapped the content IDs and related content records that had been ingested from ITV's internal platforms. The resulting unified and cleansed data, including additional complementary ITV data from MetaBroadcast's repository, was exported to ITV.

ITV is now assured that this consolidated metadata repository can be used to verify and update data from additional data sources that ITV may want to include in its overall metadata architecture. ITV has data-driven insight about metadata sources and potentially redundant metadata.

By utilising MetaBroadcast's iterative systems-based Audit, ITV is now able to ensure the use of consistent data across all lines of business, improve operational efficiency due to the use of high-quality data and identify opportunities for metadata-related cost reduction.

"As part of a wider Metadata Programme at ITV, to analyse and improve our descriptive metadata to ensure it is complete, validated and accurate, we engaged with MetaBroadcast to help us audit and cleanse descriptive data held across our internal systems.

The process which included MetaBroadcast closely working with both ITV technical and operational teams helped us understand where data was inconsistent across common content. Working with MetaBroadcast, cleansing rules were defined, resulting in a final report of improved metadata. Regular meetings and feedback sessions ensured the process went smoothly from start to finish and resulted in a really positive outcome for ITV."

- **Louise McIntosh,**
  Product Manager - Catalogue
  Management Technology