

# Documenting the Archive

— using content analysis techniques

## Внимание!

- Данный перевод **НЕ** претендует на аутентичность и может содержать отдельные неточности.
- Оригинал этого документа находится по адресу: <http://www.ebu.ch>

## Документирование Архива

— с использованием методики анализа контента

Alberto Messina  
RAI CRIT

Цель данной статьи – дать членам EBU базовую информацию о возможных преимуществах использования **методики анализа контента** для документирования телевизионных и радио архивов. Она основана на большом опыте, полученном за последние годы RAI в этой области.

В статье также анализируется влияние новых методов документирования архивов на традиционную практику работы.

В последние годы два крупных события оживили интерес вещателей к инструментам автоматического извлечения информации:

- а) феноменальный рост Интернета; и
- б) заметное падение компьютерных цен.

Первое позволило обычным пользователям пользоваться невообразимыми тоннами информации. Исследователи технологии информации и связи (ICT) сумели найти эффективные методы представления, обмена и поиска этого огромного наследия знаний. Инструменты, позволяющие достичь этих целей, варьируются от новых языков представления знаний (например, OWL) до применения уже известных или новых методов искусственного интеллекта (самые знаменитые – SVM, MLP и HMM). Среди всего этого фундаментальную роль играют **алгоритмы извлечения информации**, позволяющие извлекать неявные знания из необработанных данных (что является источником концепции “анализа контента”, изображенной на Рис. 1).

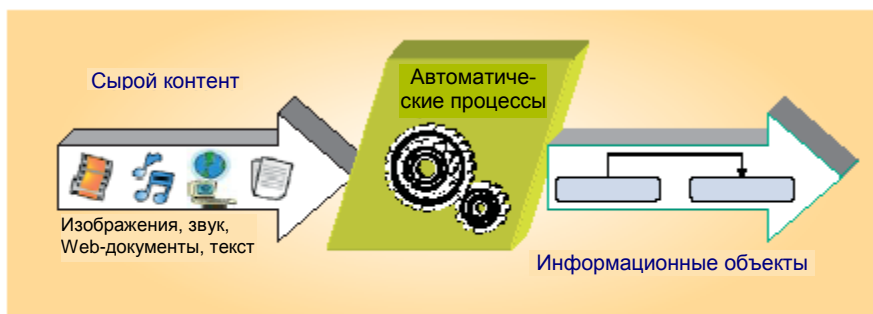


Рис. 1  
Концепция анализа контента

Обычно внедрение подобных методов влечет за собой использование существенных вычислительных ресурсов – будь то необходимый объем памяти или чистое перемалывание чисел. Особенно это касается алгоритмов, работающих с мультимедийными источниками информации, таким как видео и аудио, из-за их большого веса.

Хотя корни многих современных принципов анализа мультимедийного контента зародились несколько десятков лет тому назад, они не могли обеспечить действительно жизнеспособные решения вследствие большой вычислительной сложности. Вычислительные платформы показали свою истинную мощь только в последние десять лет.

Мощные методы представления знаний, вместе с бесподобной вычислительной мощностью, позволили возобновить исследования – сегодня мы можем с новой силой внедрять механизмы искусственного интеллекта в отраслях промышленности.

Вследствие всего этого вещатели – а кое-кто из них обладает важными историческими аудиовизуальными архивами – являются главными потребителями применения таких методов. Но как это все разумно реализовать?

## Сокращения

<b>ASR</b>	Automatic Speech Recognition Автоматическое распознавание речи
<b>HMM</b>	Hidden Markov Model Скрытая модель Маркова
<b>MLP</b>	Multi-Layer Perceptron Многоуровневое восприятие
<b>NLP</b>	Natural Language Processing Обработка естественного языка
<b>OWL</b>	Web Ontology Language ( <i>вместо</i> “WOL”) Язык веб-онтологий
<b>SVM</b>	Support Vector Machine Метод опорных векторов

## Что предлагают разные инструменты в сфере документирования архива

### С точки зрения архива

С точки зрения владельцев крупных архивов, например, RAI и некоторых других членов EBU, инструменты автоматического извлечения информации считаются большими помощниками в снижении расходов, связанных с процессом документирования архива. Идея использования мощных компьютеров для работы, обычно выполняемой квалифицированным персоналом<sup>1</sup>, представляет для этих вещателей чрезвычайно ценную возможность. Особенно это касается жанров, не требующих высокого уровня сложной аннотации для получения приемлемого уровня качества документации, например, новостей.

Однако теория и практика показывают, что мечте о замене команды документалистов бессловесными кластерами кремниевых микросхем не суждено сбыться – если не заняться этим всерьез. Характер этих машин связан с неоспоримыми фактами:

- **Проблемы обучения**

Инструменты автоматического извлечения информации необходимо “обучить” наборам входных данных, служащих “обучающими шаблонами”. Они начинают лучше работать, находя фрагменты информации, подпадающие под статистические диапазоны учебных примеров. Количество учебных этапов приближается к потенциальному количеству статистически различных категорий входного материала относительно извлекаемой информации. Кроме того, количество и свойства этих категорий могут со временем меняться, в зависимости от изменчивости выразительного языка (например, новостные программы могут часто менять свою “сценографию”, “стиль съемки” и общий “формат программы”).

- **Качество результатов**

Уровни *точности*<sup>2</sup> и *recall*<sup>3</sup> понижаются по мере увеличения семантической сложности извлекаемой информации. Парадоксально, но этот аспект концентрирует человеческие действия на подтверждении самых сложных уровней информации, где в то же время стоимость вмешательства гораздо выше, чем на нижних уровнях (из-за более высокого уровня качества).

- **Влияние на процессы**

В области системного управления необходимо провести тщательные исследования расходов, связанных с установкой, наращиванием и развитием систем, выполняющих роль ведущего узла в инструментах автоматического извлечения. Сегодня нет никакой конкретной информации о влия-

<sup>1</sup> Не предусматривая никаких социальных соображений, связанных с сохранением of employment levels.

<sup>2</sup> Здесь точность = отношение между числом «обнаружений» корректной информации и общим числом «обнаружений» информации.

<sup>3</sup> Здесь recall = отношение между числом «обнаружений» корректной информации и всей информацией, присутствующей в материале, как указывал бы идеальный детектор.

нии на архивную среду в случае внедрения таких инструментов в крупномасштабных промышленных процессах. Вследствие необходимости проверки выхода автоматических машин общая эффективность процесса документирования определенного фрагмента материала может существенно измениться, даже при своевременной оптимизации в разных местах цепи.

**История: вспомним цель**

Опыт RAI в этой области показывает, что роль **модели документации** имеет первостепенную важность, даже когда на сцену вступают автоматические процессы. Но в каком смысле это понимать?

Синтетическое (не идеальное) определение: модель документации можно определить как систему для классификации архивных элементов в закрытом наборе структур данных под названием **информационные объекты**. Информационные объекты, участвующие в модели документации, служат средством представления реальных объектов (физических объектов, а также абстрактных понятий) в конкретной информационной системе (см. Рис. 2).

Первоочередная задача любой реальной информационной системы, нацеленной на представление этого мира, заключается в обеспечении конкретных средств для **эксплуатации** находящихся в ней архивных элементов. “Эксплуатацию” можно рассматривать как действие, когда пользователь реально *пользуется* архивными элементами. С другой стороны, **пользование** – это действие, когда пользователь просто собирает информацию путем чтения/просмотра/прослушивания информации, хранящейся в системе.

Архивные элементы бывают двух классов: аудиовизуальный материал и информация. Это отличие, возможно, на первый взгляд странное, имеет ключевую важность. На самом деле, если первое представляет собой классический случай, когда пользователи архива хотят получить реальные фрагменты аудиовизуального материала, то второе относится к тому случаю, когда информация является принципиальной целью взаимодействия пользователей с архивом. Примеры можно найти в ситуациях, когда пользователь лишь интересуется, произошло ли определенное событие, или читает детали описания места и т.д. В этой интерпретации традиционное понятие **метаданных** открывает два нюанса: в процессе поиска фрагмент “метаданных” является архивным элементом, который служит информативной связью с другим элементом. Возможны четыре случая:



Рис. 2  
Представление объектов реального мира с информационными объектами

- Поиск аудиовизуального элемента с помощью информации**  
 Этот случай представляет эксплуатацию архива *per antonomasia*. Если стартовой точкой является спецификация ограничений, которым должна соответствовать информация (например, предполагаемые значения или диапазоны атрибутов, достоверность определенных связей), то финишной точкой будет фактический поиск аудиовизуального материала, для которого действительно указаны ограничения. Это традиционный метод использования информации в качестве "метаданных" для поиска нужного A/V материала.
- Поиск информации с помощью аудиовизуального элемента**  
 Это более сложный способ доступа к архивной информации. Здесь сам аудиовизуальный материал является носителем информации, которая нужна пользователю. Следует заметить, что конечной целью эксплуатации в данном случае является только информация. Стартовой точки (материала-носителя) можно достичь в любом из четырех упомянутых здесь случаев.
- Поиск информации с помощью информации**  
 В этом сценарии задача пользователя опять состоит в поиске информации. Но здесь это достигается с помощью других фрагментов информации, служащих "метаданными" для целевой информации. Типичные примеры: поиск тем, содержащихся в речи определенного персонажа, поиск подробностей определенного исторического события.
- Поиск аудиовизуального элемента с помощью аудиовизуального элемента**  
 Эта категория включает случаи, когда аудиовизуальный материал ищется по подобию на основе исключительно аудиовизуального контента, т.е. независимо от выраженного значения и основной семантики. Типичные инструменты здесь – это запросы на примерах на основе контента с использованием индексов аудиовизуальных свойств.

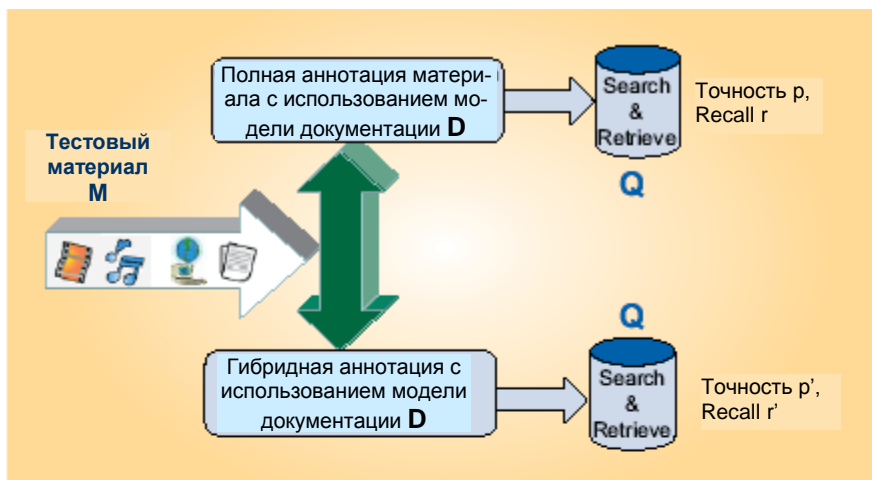


Рис. 3 Сравнение параметров качества поиска

Главная концепция этих сценариев – что инструменты извлечения информации, основанные на автоматическом анализе контента, должны оцениваться на основе их способности обеспечивать экземпляры данных с контролируемой степенью точности.

Точность recall, в свою очередь, следует оценивать сравнением уровней "точности" и "recall", полученных методами анализа контента и с традиционной ручной аннотацией. Разумеется, сравнение должно проводиться без изменения модели документации или справочного материала (учебного и тестового).



Рис. 4 Этапы процесса автоматического извлечения информации

Другими словами, данную модель оценки можно считать тестовой системой, в которой, во-первых, полная ручная аннотация производится с помощью определенной модели документации D на наборе материалов M. На втором этапе, с теми же D и M, документирование производится идентично гибридной системой документации, имеющей инструменты автоматического извлечения. Делается прямое сравнение – из обеих систем запрашивается один и тот же набор запросов Q. Затем оце-

нивается точность и recall списка элементов, с опорой на идеальные результаты (известные *a priori*).

Это проиллюстрировано на *Рис. 3*.

### **Анализ соответствующих проблем**

В этом разделе содержится краткий обзор вопросов, связанных с использованием инструментов автоматического извлечения информации.

На *Рис. 4* показана типичная последовательность интеграции инструмента автоматического извлечения информации. Первый этап, сопровождающий этап выбора тестового материала, часто называют **генерированием свойств**, т.е. из выбранного материала извлекается “разумный набор” свойств. Обычно это делается по типу материала (т.е. аудио или видео) и по признакам и индикаторам, найденным в технической и научной литературе по классу проблем, для которых предназначен данный инструмент.

На этапе выбора свойств отбираются самые обещающие выборки из первого этапа. В общих словах, сложность этой задачи зависит от характеристик “отделимости”, проявляемых классами идентифицируемой информации относительно выбранных поднаборов свойств.

Тривиальный пример: если задача классификатора состоит в отделении черных кадров видеоряда от нечерных, то наиболее вероятно, что самым обещающим свойством станет отслеживание усредненного уровня яркости в видео кадре.

В менее банальных случаях все гораздо сложнее. Этап обучения и подстройки во многом зависит от вида инструмента. В основном он занимается оптимизацией функций для получения стабильного набора параметров используемых классификаторов. На этапе тестирования классификатор тестируется по шаблонам, которых не было на этапе обучения, с целью оценки общей производительности системы. Наконец, если все сходится, начинается этап работы. Однако с этим процессом связаны очень серьезные проблемы, а именно:

- **Высокие требования к разработке данных**

Обычно вещательные архивы содержат тысячи часов материала, а темпы роста активного архива могут достигать нескольких десятков часов в сутки. Это ведет к двум проблемам: с одной стороны, большой объем материала означает большое количество тестовых шаблонов, а с другой стороны, быстро вводит новые типы и парадигмы. Это влияет на стабильность обучающих параметров и количество необходимых классификаторов. Весьма вероятно, что понадобится частое переобучение и что надо будет добавлять новые классификаторы. Это может повлечь за собой переработку архивного материала.

- **Изменчивость операционной сферы**

Обучение классификаторов обычно дает лучшие результаты, когда классы, которые надо идентифицировать и связать с опорными шаблонами, имеют высокий коэффициент отделимости, т.е. когда есть комбинация свойств контента, позволяющая определение четких классификационных границ. Вследствие многообразия операционных требований вещательного архива материал редко классифицируется абсолютно, раз и навсегда (например, фрагмент материала может классифицироваться под разными категориями в зависимости от планируемого использования, а не просто по характеристикам контента).

- **Несоответствие области применения**

Технику анализа контента и извлечения информации можно представить в контексте сфер применения, задачи которых могут быть далеки от требований процесса документирования вещательного архива. Например, многие современные методы автоматической классификации текста лучше подходят для работы со сформированными текстовыми источниками, чем с транскрипциями речевого языка (ASR). Проблемы несоответствия областей серьезно влияют на этап обучения и подстройки алгоритмов. Большую часть знаний данной области приходится привязывать к существующим инструментам, не зная *a priori* о влиянии этого введения на производительность инструмента.

- **Несовпадение детализации информации**

Современные инструменты извлечения информации на основе анализа контента обычно работают с фрагментом входящего материала и выдают набор метаданных с четкой структурой и семантикой. Детализация связи извлеченных данных с входящим материалом не может управляться пользователем (если только он сам не разрабатывал этот инструмент, но это случается редко). Например, инструмент классификации по тексту на основе NLP может связывать ярлык класса (например, "Спорт") с фрагментом материала целиком, в то время как пользователь мо-

жет предпочесть динамическую классификацию, меняющуюся в течение материала. Кроме того, некоторые классификаторы могут ограничиваться встроенными схемами классификации, которые трудно (или невозможно) использовать со старыми или пользовательскими схемами.

## Разумный метод

Недавно RAI нашла полезным отказаться от “естественного” подхода к технике анализа контента. Проще говоря, вместо того, чтобы интегрировать существующие инструменты в среде документирования и обеспечить компоненты интеграции в модель документации, лучше начать с целевой информационной модели, заполнив ее, а потом, *в обратном порядке*, определить, какие есть инструменты и с каким уровнем точности они дают результаты. Такой анализ можно назвать “анализом функции процесса”. Начиная со спецификации модели документации, анализ функции процесса идентифицирует элементарные функции, необходимые для заполнения экземпляра этой модели. Пример результатов анализа функции процесса показан в *Таблице 1*.

**Таблица 1**  
**Пример результатов анализа функции процесса**

Область	Что		Как	
	Сфера	Объект	Источник	Результирующая функция
Контент	Описания	Лицо	Видео	Распознавание лиц
		Лицо	Аудио	Распознавание речи
		Место	Видео	Распознавание места
		Физический объект	Видео	Распознавание объектов, отслеживание объектов
		Информация движения камеры	Видео	Распознавание движения камеры
	Темы	Лицо	Текст	Извлечение названного объекта
	Классификации	Категория	Текст	Семантический анализ текста
	Классификации	Категория	Видео	Классификация места (например, на улице / в помещении)
	Описания	Тишина	Аудио	Распознавание тишины
	Текст	Текст	Аудио	Автоматическая транскрипция речи
Идентификация	Вклад	Лицо	Видео	Транскрипция видео в текст
	-	Награды	Текст	Web-индексирование, анализ web-данных
Редакторские части	-	Сцена	Видео и аудио	Обнаружение и сегментация сцен
		Новостной элемент	Аудио	Распознавание и группирование говорящих
		Спортивный элемент	Аудио и видео	Обнаружение главных событий

Ключевая концепция в данном случае – изоляция элементарных функций процесса документации (каждая из которых представлена в таблице строкой) на основе структуры модели документации (представленной в колонке “Что”) и связывание функции инструмента с каждой элементарной функцией (представленной в колонке “Как”). Например, если требуется дать описание контента через описание лиц, участвующих в сцене, то это можно сделать с помощью инструментов распознавания лиц и/или речи на основе анализа мультимедийного контента. Если надо идентифицировать редакторские части, например, сцены в художественном фильме, потребуется алгоритм обнаружения и сегментации сцен.

Эта стратегия дает некоторые преимущества, которые можно резюмировать следующим образом:

- Оптимизация интеграции инструментов. Оцениваются и интегрируются только те инструменты, которые могут обеспечить функции, необходимые для требуемой информации.
- Оптимизация качества извлеченной информации. Качество можно оптимизировать только с помощью самых лучших инструментов для обеспечения функции определенного процесса.

Этап выбора инструментов должен учитывать, с одной стороны, доступный на сегодня уровень “точности” и “recall”. С другой стороны, в промышленном контексте следует учитывать объем работ, необходимый для интеграции существующих инструментов в инфраструктуру документации (например, в смысле механизмов обновления, параметров конфигурации, условий выполнения и т.д.). Важно отметить, что общее решение – использовать ли автоматический механизм, выполняющий одну из элементарных

функций – должно подкрепляться взвешенной комбинацией обоих аспектов: следует отказаться от инструментов с большой “точностью” и “recall”, но недоработанной программной архитектурой, как и от инструментов с отличными программными характеристиками, но дающих плохие результаты.

## Краткая текущая классификация

Полный анализ текущей ситуации не входит в рамки данной статьи, дающей главным образом введение в предмет. Заинтересованный читатель может найти хороший аналитический обзор в [1], а глубокий трактат о фундаментальной теории – в [5]. Вместе с тем, в контексте данной статьи можно рассмотреть инструменты автоматического извлечения информации на основе контента по следующим категориям:

- **Методы представления на базе свойств нижнего уровня (LL)**

К этому классу принадлежат все алгоритмы, чья цель – извлечение базовых свойств из аудиовизуального контента и представление их в структурах данных, подходящих для прямого поиска на основе определенного подобия. В случае видео эти свойства обычно относятся к **цвету** (например, цветовые гистограммы), **текстуре** (например, векторы текстурной сигнатуры, состоящие из измерений контраста, зернистости и направленности), **форме** (например, векторы кривизны и ориентации) и **движению** (например, поля векторов движения). В случае аудио типичные извлеченные свойства относятся к **частоте** (например, средние частоты, полоса частот, спектральные коэффициенты) и **времени** (например, частота переходов через нуль, кратковременная энергия). Этот уровень техники основан на предположении, что базовые свойства, например, цвет, форма, текстура, движение и частота являются измерениями, по которым человеческое восприятие классифицирует изображение, звук и видеоряд перед семантической интерпретацией. Оценивать качество инструментов LL следует согласно их точности в представлении явных перцепционных аспектов аудиовизуального материала.

- **Методы структурирования медиа на базе формальных свойств (FS)**

Эти методы строятся на результатах LL с целью извлечения структурной информации из аудиовизуального контента. Классическая проблема, которую пытается решить FS – это проблема “обнаружения планов”, т.е. восстановление *a posteriori* последовательности коммутации датчиков, визуальных или слуховых, которая была в ходе записи аудиовизуального материала. Частый случай обнаружения планов – проблема “индексации видео”, т.е. создание сжатых структур временной и визуальной информации, учитывающей важные аспекты видеоряда, позволяя нелинейное, адаптивное к контенту согласование материала. В аудио аналогом является проблема “классификации аудио сегментов”, т.е. разделение аудио сегментов на тишину, речь, музыку, шум.

- **Методы обнаружения и идентификации объектов (OD)**

Этот класс методов опирается на LL и частично на FS, их цель – решение проблем, связанных с обнаружением и идентификацией объектов, по-разному связанных с аудиовизуальным контентом. Оценка соответствия объектов сильно зависит от типа целевого описания, который, в свою очередь, связан с типом применения результата методов OD, а также с характером и структурой выбранной информационной модели. Самые частые проблемы в методах OD – распознавание лиц и людей, распознавание и отслеживание физических объектов.

- **Методы классификации контента (CC)**

Классификация контента считается одной из основных функций автоматической обработки аудиовизуального контента. Типичный подход основан на выборе правильной вероятностной модели, способной максимизировать вероятность набора данных свойств нижнего уровня, взятого за обучающий набор, а затем использовать такие модели для классификации новых элементов. Методы классификации могут также учитывать информацию высшего уровня, например, OD или MS, для повышения качества вероятностных моделей. Текстовые методы выполняют задачу классификации с помощью статистических или лингвистических моделей, обычно после автоматической транскрипции речи в текст. Еще один аспект, тесно связанный с классификацией – это “маркировка классов”, т.е. ассоциация ключевых слов с определенными классами как первая попытка сопоставления классификаций по свойствам с реальным значением, выраженным аудиовизуальным контентом.

- **Методы структурирования, управляемые семантикой (SS)**

В самых общих чертах, методы семантического извлечения в основном строятся поверх других методов (LL, FS, OD и CC) и решают оставшуюся проблему заполнения пробела между инфор-

мацией нижнего уровня (обычно LL, FS) и семантическим контентом аудиовизуального материала, как выражается человеком при восприятии контента. Точнее, методы структурирования, управляемые семантикой (SS) нацелены на идентификацию редакторских компонентов аудиовизуального произведения, т.е. составных частей, задуманных создателем. Как сказано в литературе [5], эта задача в основном зависит от жанра контента, вследствие логичного динамизма и изменчивости, характерных для телевизионного языка, например, что касается таких измерений как целевая аудитория, программные форматы и цель. По этим причинам проблема идентификации сцен в фильме может быть совершенно иной в смысле автоматических методов и успешного использования инструментов, чем проблема идентификации новостных, спортивных элементов или развлекательных программ.

Как и в любой системе классификации, вышесказанное представляет только точку зрения автора. Другие мнения по этому вопросу представлены в [2] и [3].

## Оценка влияния на процессы

На основании существующей практики и научно-технического прогресса можно сделать вывод, что инструменты автоматического документирования в среднем являются достойными кандидатами на внедрение в процессе документации архива. Это верно, но не надо забывать, что всегда в некоторой форме необходим человеческий контроль (например, обучение, обновление, подгонка результатов).

Как оценить возможность принятия или непринятия подобных инструментов? Возможный ответ – сравнение качества извлеченной информации с информацией, извлеченной человеком, при постоянной частоте ошибок.

Проводя это сравнение, мы можем достичь компромисса, выраженного в математической форме:

$$\sum_{i=1}^N C_i^E(\varepsilon_i) + C_i^{CHK}(\varepsilon_i) < \sum_{i=1}^N C_i^H(\varepsilon_i) \quad (1)$$

... где  $N$  – число идентифицированных функций процесса (т.е. строки в *Таблице 1*), для которых оценивается использование автоматических методов,  $\varepsilon$  – принятая ошибка поиска (например, принятое F-измерение),  $C^E$  – стоимость разработки,  $C^{CHK}$  – стоимость в контрольной точке, а  $C^H$  – стоимость полностью человеческой документацией.

Обратите внимание, что расходы, связанные с автоматическими и с человеческими процессами, имеют разный характер. Поэтому для того, чтобы сравнение имело смысл, их надо выразить в эквивалентной единице измерения. Однако, если в человеческих расходах проще опираться на параметры стандартной единицы (например, рабочей силы), то для автоматических компонентов такая оценка сложновата. Стоимость амортизации оборудования, лицензии на программное обеспечение, текущие расходы на аппаратные средства и программное обеспечение и наращивание инфраструктуры могут быть важными элементами в выработке конкретных указаний в данном направлении.

*Уравнения (1)* вполне достаточно для принятия решений, предполагая, что внедрение автоматических машин ограничится соответствующими функциями, для которых предназначена данная машина (например, автоматический перевод речи в текст vs. ручная транскрипция). Однако в оценке влияния необходимо учитывать и возможные (положительные) синергии с процессами и функциями, которые остаются человеческими (например, внедрение алгоритма распознавания лиц может снизить затраты на аналитическую идентификацию участников программ, выполняемую документалистом).

Эти соображения в синтетическом смысле означают, что возможность использования автоматических инструментов удобна, если стоимость автоматической работы плюс человеческого контроля ниже стоимости человеческой работы при постоянном уровне точности/recall в архивном поиске. Это условие надо оценить и усреднить по общему количеству функций процессов, в которых планируются инструменты анализа контента.





Alberto Messina работает в Центре исследований и технологических инноваций *Radiotelevisione Italiana* в Турине.

Начал сотрудничать с RAI, разработав собственный тезис объективной оценки качества видео кодирования MPEG-2. Окончив в 1996 г. университет, работал в RAI инженером, начав свою карьеру среди разработчиков мультимедийного каталога RAI. Его сегодняшние интересы – от изучения инструментов и методов обмена и представления знаний в вещательной среде (в форме файловых форматов, стандартов метаданных и систем управления контентом) ... до сферы анализа контента и алгоритмов извлечения информации. Недавно он начал исследования семантического извлечения информации из анализа аудиовизуального материала.

Участвуя в нескольких других междисциплинарных проектах RAI, г-н Мессина сотрудничал в проектной группе EBU P/META и является активным членом проектов P/TVFILE, P/MAG и P/CP. Также работает сейчас в проекте EU PrestoSpace в области доступа к метаданным и их распространения.

## В заключение

Члены EBU, владеющие важными архивами, могут оказаться в положении оценки использования инструментов автоматического извлечения информации на основе анализа контента в области документации архива. Возрождение интереса побуждает заимствовать методы из промышленности, в сочетании с двумя факторами: наличием огромной вычислительной мощности по скромной цене и параллельным наличием передовых инструментов представления знаний.

Вещательные архивы представляют явный сценарий применения, но, как показано в настоящей статье, необходим тщательный анализ по строгим принципам, чтобы не расточать драгоценные ресурсы на исследование того, что может оказаться бесполезной (хоть и прелестной) игрушкой.

В частности, опыт RAI в этой области подсказывает, что следует сохранить модели документации как основную точку принятия решений по выбору и интеграции инструментов в процесс документации. Кроме того, вещателям необходимо провести обширные исследования для оценки влияния такого введения в смысле дифференциальной стоимости.

## Ссылки

- [1] W. Bailer, F. Holler, A. Messina, D. Airola, P. Schallauer and M. Hausenblas: **State of the art of Content Analysis Tools for Video, Audio and Speech**  
Deliverable 15.3 of the IST PrestoSpace project.
- [2] C.G.M. Snoek and M. Worring: **Multimodal Video Indexing: A Review of the State-of-the-art**  
Multimedia Tools and Applications, 25 (1): pp 5 - 35, January 2005.
- [3] M. Roach, J. Mason, L.-Q. Xu and F.W.M. Stentiford: **Recent trends in video analysis: a taxonomy of video classification problems**  
6th IASTED Int. Conf. on Internet and Multimedia Systems and Applications, Hawaii, Aug 12 - 14, 2002.
- [4] A. Messina and D. Airola Gnota: **Automatic Archive Documentation Based on Content Analysis**  
IBC 2005 Conference Publication.
- [5] S. Theodoridis and K. Koutroumbas: **Pattern recognition – second edition**  
Academic Press, 2003.