DOI: https://doi.org/10.37075/ISA.2026.3.02

Принципи за автоматизирано структуриране на метаданни в платформи за видео споделяне

Явор Табов*

Резюме: Бързото развитие на дигиталните технологии и нарастващото значение на онлайн видео съдържанието трансформират начина, по който информацията се създава, разпространява и потребява. Платформите за видео споделяне се утвърждават като ключови среди за комуникация и обмен на знания, което поставя високи изисквания към управлението и откриването на съдържание. В този контекст метаданните играят съществена роля за структуриране и оптимизация на видео ресурсите. Настоящата статия разглежда принципите за автоматизирано структуриране на метаданни, като анализира тяхната същност, видове и функции в уеб среда. Изследването обхваща тенденциите в развитието на видео платформите и технологичните аспекти, свързани с видео кодирането, методите за компресия на видео и компонентите на потребителския интерфейс. Основен акцент е поставен върху автоматичното генериране на метаданни чрез мултимодален анализ, който интегрира аудио, визуална и текстова информация. В заключението се подчертава значението на автоматизираното структуриране на метаданни за функционирането и развитието на платформите за споделяне на видео. Отбелязва се ролята на мултимодалните подходи за по-точно описание и интерпретация на видеата с цел подобряване на процесите на индексиране, търсене и персонализация, и създаване на условия за по-ефективно управление и предоставяне на видео съдържание.

Ключови думи: метаданни, платформи за видео споделяне, мултимодална интеграция

JEL: C88, L86.

Увод

Развитието на дигиталните технологии през последните десетилетия доведе до дълбока трансформация в начина, по който се създава, разпространява и потребява аудиовизуално съдържание. Платформите за видео споделяне се утвърдиха като ключова част от съвременната интернет среда, като предоставят възможности не само за публикуване и гледане

* Явор Табов е доктор, главен асистент в катедра „Информационни технологии и комуникации“ на УНСС.

на видеа, но и за активно социално взаимодействие, персонализирано съдържание и мащабно разпространение на информация. С нарастването на обема от видео материали възниква необходимост от по-ефективни механизми за тяхното организиране, откриване и управление, което поставя метаданните в центъра на технологичната и методологичната рамка на тези платформи.

В този контекст настоящата статия разглежда принципите на автоматизираното структуриране на метаданни в платформи за видео споделяне, като проследява тяхната роля както в описанието и класифицирането на видео съдържанието, така и в подпомагането на процеси като търсене, препоръки, достъпност и оптимизация на потребителското изживяване. Специално внимание е отделено на същността на метаданните като средство за структуриране и обмен на информация, на техните основни видове и функции в уеб базираните системи, както и на значението им за ефективното управление на мултимедийни ресурси.

Статията представя също особеностите на съвременните платформи за видео споделяне и основните тенденции в тяхното развитие, включително нарастващата роля на интерактивността, персонализацията и алгоритмичните системи за препоръки. Наред с това се акцентира върху технологичната основа на тези среди, свързана с видео кодирането, форматите за компресия, адаптивния стрийминг и ключовите компоненти на потребителския интерфейс, които определят начина, по който съдържанието достига до аудиторията.

Основен акцент в изследването е поставен върху автоматизираното генериране и структуриране на метаданни чрез съвременни мултимодални подходи, които обединяват аудио анализ, визуален анализ и обработка на естествен език. Подчертава се, че именно интеграцията на тези методи създава устойчива и мащабируема основа за по-точно описание на видео съдържанието и за развитието на по-ефективни системи за индексиране, търсене и персонализирано предоставяне на информация. По този начин статията откроява значението на автоматизацията като ключов фактор за управлението на нарастващата сложност в съвременните видео платформи.

1. Метаданните като инструмент за структуриране и обмен на информация

Метаданните се отнасят до информация, която описва други данни, като например видеофайл, аудиозапис или уеб страница. Те играят ключова роля за подобряване на откриваемостта на видео съдържанието както за отделни потребители, така и за организации, които публикуват видеа на своите уебсайтове или на платформи като Google, iTunes, YouTube, MySpace, Youku и Vimeo. Метаданните не променят съдържанието на видеото, а служат като информативен етикет. Когато метаданните се създават или извличат от видео, те позволяват избора на подходящи ключови думи и описания, които подпомагат оптимизацията за търсачки (SEO). Търсачките използват метаданните, за да определят позиционирането на съдържанието в

резултатите от търсене и да генерират кратки описания, показвани в тези резултати. Този процес подобрява както видимостта, така и качеството на трафика към уебсайтовете. Метаданните са особено важни за видео материали, тъй като видеофайловете нямат присъщи текстови елементи, които могат да се използват за търсене по ключови думи. Въпреки че създателите на видеа могат да включват ключови думи в имената на файловете, е необходима допълнителна описателна информация, за да могат компютрите и търсачките точно да интерпретират и откриват видео съдържание (Sathish Kumar и др., 2019).

В повечето стандарти за метаданни няма изричен механизъм за представяне на множество изгледи на една и съща сцена, особено когато тези изгледи се различават по времетраене. За разлика от това, поддръжката на многоканално аудио обикновено е поразвита поради по-дългата история на стандартизация (Bailer и др., 2009).

По принцип метаданните означават „данни за данни“, тоест те предоставят информация, описваща природата и характеристиките на други данни. Те могат да се сравнят с индекса на книга, който съдържа кратка информация, обобщаваща съдържанието ѝ. Чрез преглед на индекса читателят може да добие представа за книгата, без да чете целия текст, въпреки че самият индекс не включва подробна информация. По подобен начин метаданните предоставят структурирано обобщение, което подпомага организацията и извличането на данни. В контекста на търсачките метаданните играят съществена роля при намирането на конкретни уеб страници, които съответстват на заявката на потребителя, като по този начин подобряват точността и ефективността на извличането на информация (Sathish Kumar и др., 2019).

Метаданните представляват данни, които предоставят информация за други данни. Популярни видео платформи като YouTube, Twitch, Instagram, TikTok и Netflix функционират като уеб базирани приложения, изградени върху стандартни уеб технологии. Като такива системи те разчитат на структурирани метаданни, за да описват, управляват и предоставят видео съдържание ефективно. Метаданните играят ключова роля във формирането на методологичните аспекти на структурирането на съдържанието, като подпомагат ефективното му откриване, оптимизацията на стрийминга и взаимодействието с потребителите чрез определяне на начина, по който видео ресурсите се представят и обработват в уеб среда. Използването на динамично структурирани конфигурационни формати в комбинация с бази данни позволява гъвкаво представяне и персонализация на данните, без да се налагат постоянни модификации в основния кодов блок на дигиталните платформи (Милев, 2025а). Структурирането на съдържанието на тези платформи в голяма степен разчита на различни видове метаданни, които определят как съдържанието се описва, организира, достъпва и доставя. В уеб базирани приложения, реализирани чрез HTML, се използват различни HTML елементи и мета елементи за представяне и структуриране на метаданни.

За по-ясно разбиране на понятието метаданни, те могат да бъдат разделени на няколко категории, които отразяват различните им роли в информационните системи. Класификацията на метаданните зависи от конкретната област, изискванията на потребителите и естеството на представяните ресурси или обекти (Formenton и др., 2022). В научната литература метаданните могат да бъдат класифицирани в следните типове: описателни, структурни, административни и специфични за медии.

Описателните метаданни предоставят информация за съдържанието и предназначението на уеб страница, като помагат на търсачките и потребителите да разберат за какво се отнася тя, и играят централна роля при откриването на съдържание и оптимизацията за търсачки. Тук обикновено се включват следните HTML елементи:

- HTML заглавие (title tag) – определя заглавието на страницата, което се показва в резултатите от търсене и в разделите на браузъра, като пряко влияе върху честотата на кликване. HTML заглавията винаги са били и остават най-важният сигнал, който търсачките използват, за да разберат съдържанието на страницата. Лошите заглавия на уеб страниците са като лоши заглавия на книги в горния пример. Всъщност, ако HTML заглавията се считат за неподходящи или неописателни, търсачките могат да ги променят (Patel и др., 2018).
- Мета описание (meta description) – предоставя кратко обобщение на страницата, което помага на потребителите да преценят нейната релевантност към тяхното търсене (Sharma и др., 2024). Тази информация също влияе върху ангажираността и честотата на кликване, въпреки че търсачките могат да я модифицират в зависимост от заявката (Patel и др., 2018).
- Мета ключови думи – съдържат термини, свързани със съдържанието на страницата; въпреки че значението им е намаляло, те все още предоставят контекстуални сигнали (Sharma и др., 2024).

Структурните метаданни определят как различните части на даден ресурс са организирани и свързани. Те включват елементи като уникални идентификатори, номера на страници и структурни характеристики като съдържание и индекси. Освен това описват връзките между набори от данни и подпомагат свързването между отделните компоненти на дигитален ресурс (Mosha и др., 2023). Тази категория обикновено включва следните HTML елементи:

- Карта на сайта (sitemap) – помага на търсачките систематично да обхождат и индексират всички страници на уебсайт, като предоставят информация за йерархията на страниците, честотата на обновяване и относителния приоритет, което подобрява откриваемостта и оптимизацията на сайта.
- Тагове за връзки (link tag) – дефинират връзките между документи и външни ресурси, като стилове, алтернативни версии или канонични URL адреси (Sharma и др., 2024).

Административните метаданни описват разрешенията, правата, съхранението и използването на данните, като подпомагат правилното управление на съдържанието, съответствието с правни и лицензионни изисквания и дългосрочната достъпност в уеб приложения (Allen, 2020). Често използваните HTML елементи включват:

- Заглавни полета на HTTP (HTTP headers) – съдържат данни, които описват техническите характеристики на уеб ресурса и начина на неговото предаване и обработване. Чрез тях се предоставя информация за типа на съдържанието, кеширането, сигурността, удостоверяването и други параметри, свързани с достъпа до и предаването на уеб съдържание (Gadient и др., 2021).
- Елементи за автор и авторско право (author, copyright) предоставят информация за собствеността, авторството и лицензионните условия.

Специфичните за медии метаданни се отнасят до описателна информация, свързана с мултимедийни елементи като изображения, видеа и PDF файлове, които са широко използвани в съвременните уеб приложения, като подобряват видимостта, споделянето и достъпността на съдържанието. Към тях спадат както метаданни, свързани с представянето на съдържанието в уеб среда, така и технически характеристики на самите мултимедийни обекти, като аудио характеристики (битрейт, честота на дискретизация), видео характеристики (резолюция, честота на кадрите) и характеристики на изображенията (размери, дълбочина на цветовете) (Sharma и др., 2024). Често използвани са следните HTML елементи и атрибути:

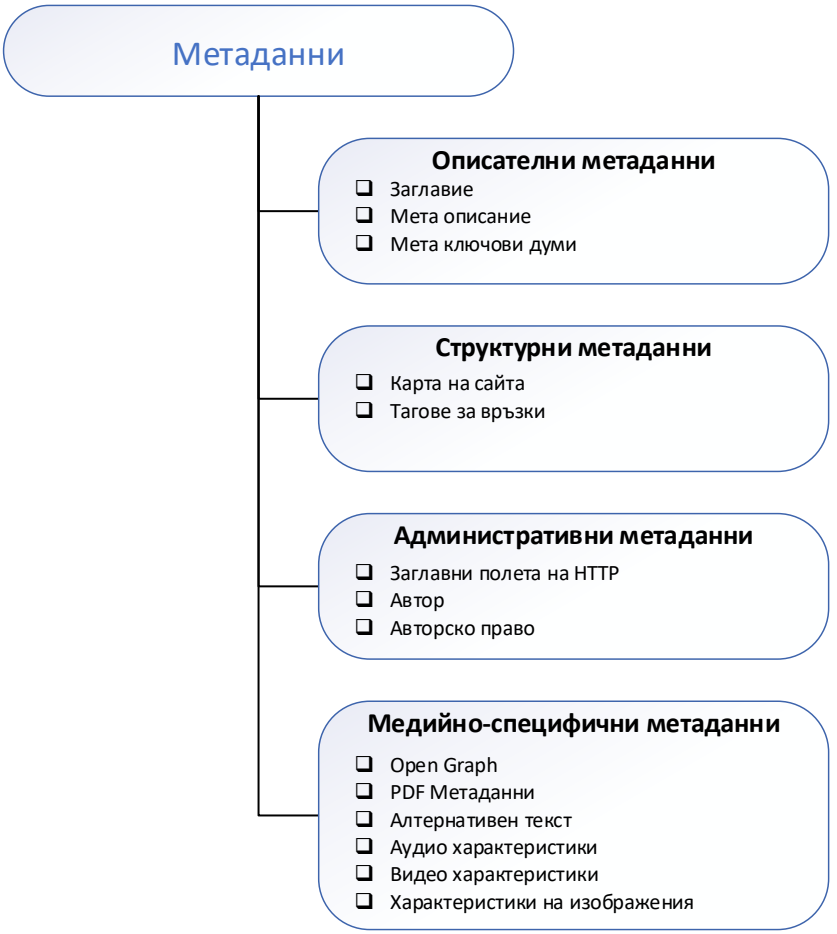
- Тагове Open Graph (OG) – предоставят структурирана информация за уеб страниците и определят начина, по който съдържанието се визуализира при споделяне в социалните мрежи. Те дефинират ключови атрибути като заглавие (og:title), описание (og:description), изображение (og:image) и адрес на ресурса (og:url), които формират т.нар. социални карти (social cards) при споделяне на съдържание (Jones и др., 2021).
- Метаданни на PDF файлове (PDF metadata) – подпомагат организацията и свързването на документни ресурси.
- Алтернативен текст (alt text) – подобрява достъпността за потребители със зрителни увреждания и подобрява индексирането от търсачките (Sharma и др., 2024).

Фигура 1 представя гореописаните елементи на метаданни, разпределени в четирите категории в контекста на уеб базирани платформи за споделяне на видео.

Метаданните предоставят подробна информация за съдържанието на видео библиотеки и подпомагат организацията и управлението на такива колекции. В контекста на дигиталните видео файлове, тези данни могат да бъдат външни – съхранявани отделно от файловете и предоставяни от дигиталните платформи – или вградени. Вградените метаданни се съхраняват като част от структурата на видео файла. Те включват описателна информация, като например данни за

авторско право, автор, ключови думи, размер на файла, медиен формат, участници, местоположение и информация за субтитри. Тези данни остават свързани с файла при копиране, споделяне или публикуване. Относно създаването си, вградените видео метаданни се разделят на два типа:

- Автоматични метаданни – генерирани от устройства или софтуер като камери. Те записват технически детайли като дата, час, модел на камерата, диафрагма, скорост на затвора, GPS локация и други настройки, обикновено във формат EXIF. Повечето EXIF данни не могат да бъдат променяни след запис.
- Ръчни метаданни – създадени от хора за описание на съдържанието на видеото. Те включват заглавия, описания и транскрипции, които улесняват търсенето и категоризацията. Ръчно добавените метаданни подобряват видимостта в търсачките и



Фигура 1. *Категории метаданни и техните елементи*
Източник: *Авторово изображение*

подпомагат маркетинга, като помагат на зрителите да откриват релевантно съдържание.

И двата типа метаданни са съществени за споделянето и откриването на видеа. Платформи като YouTube разчитат на метаданни, за да класифицират видеата и да подобрят функциите за търсене и препоръки (Sathish Kumar и др., 2019).

В обобщение, видео метаданните представляват структурирана информация, вградена в дигитален видео файл, която описва съдържанието и техническите характеристики на видеото. Те могат да бъдат извлечени от различни части на йерархичната структура на файла, като заглавни блокове или специфични сегменти от данни. Метаданните включват детайли за видео, аудио и субтитри и са от съществено значение за организиране, анализ и верификация на видео файлове в различни стандарти за кодиране (Xiang и др., 2021). Освен това, метаданните могат да бъдат външни – съхранявани отделно и предоставяни от платформите, в които се публикува съдържанието.

Метаданните са от ключово значение за уеб базираните видео платформи, тъй като позволяват ефективна организация, откриване и предоставяне на съдържание. Чрез използването на описателни, структурни, административни и специфични за медии метаданни, платформите могат да подобрят търсенето, да оптимизират стрийминга, да повишат достъпността и да подпомогнат ангажираността на потребителите, като същевременно осигуряват правилно управление, собственост и техническа ефективност на видео ресурсите.

Изградената методологична и понятийна рамка на метаданните показва фундаменталната им роля за организирането и структурирането на дигиталната информация. За да се разбере обаче тяхната практическа приложимост и значение в реална среда, е необходимо тези принципи да бъдат разгледани през призмата на съвременните платформи за видео споделяне. Именно специфичната динамика, бързото развитие и променящите се потребителски нагласи в тези платформи поставят нови изисквания към управлението на метаданните и налагат тяхната постоянна адаптация.

2. Същност и тенденции на съвременните платформи за видео споделяне

Платформите за видео представляват динамични дигитални среди, които подпомагат създаването, разпространението и взаимодействието с съдържание в голям мащаб. Платформи като YouTube, TikTok и Twitch се развиват бързо, привличайки милиони потребители и превръщайки се в основни пространства за комуникация, забавление и обмен на информация. Тяхната същност се определя не само от публикуването на видеа, но и от силно изразените социални и интерактивни функции (Bartolome и др., 2023).

Платформи като Netflix и Disney+ също претърпяват бърз растеж, като прогнозите сочат над 1,5 милиарда абонати и 18,2% пазарно проникване до 2025 г. Този ръст отразява увеличаването на гледането на видео онлайн

чрез устройства, свързани в Интернет, което променя начина, по който екранните медии се създават, разпространяват и консумират. Услугите за поточно предаване на видео не само трансформират икономическите структури на традиционното телевизионно излъчване, но и фундаментално променят моделите на медийно потребление. Зрителите вече имат достъп до съдържание при поискване, персонализирано и достъпно на множество устройства, което контрастира с фиксираните програми и линейното излъчване на традиционната телевизия. Тази промяна подчертава трансформиращото въздействие на платформите за поточно предаване на видео върху поведението на аудиторията и медийната екосистема (Evens и др., 2024).

В глобален план, видео платформите бързо се развиха в големи социални и медийни екосистеми, всяка със свои специфични характеристики, потребителски бази и изследователски интереси. Платформата YouTube, създадена през 2005 г. и придобита от Google през 2008 г., се превърна в най-голямата и влиятелна, привличайки милиарди гледания и милиони създатели на съдържание по целия свят. Тя съчетава съхранение и предоставяне на видео съдържание със силни социални и интерактивни функции, оформяйки комуникацията, забавлението и обмена на информация сред разнообразни аудитории.

Освен YouTube, някои платформи се фокусират върху специализирано съдържание и взаимодействие в реално време; пример за такава платформа е Twitch. Стартирана през 2011 г., Twitch е платформа за поточно предаване на видео на живо, която през 2022 г. е отчела 140 милиона месечно активни потребители. Основното й съдържание се състои от играчи на компютърни игри (стриймъри), които излъчват видео и аудио на живо, съпроводено с интерактивен текстов чат. Изследванията върху Twitch акцентират върху техниките за стрийминг на живо, взаимодействието в реално време и динамиката на общностите, като подчертават социалните и интерактивните аспекти на платформата (Bartolome и др., 2023).

Платформите за поточно предаване на видео предлагат по-интерактивно изживяване при гледане, като дават на потребителите по-голям контрол върху това какво гледат, кога го гледат (time-shifting) и къде го гледат (place-shifting). За разлика от традиционната телевизия, аудиовизуалното потребление вече не е обвързано с фиксирани програми. Функциите на платформите, които засилват интерактивността, дават на потребителите усещане за контрол върху техните решения и действия, като подобряват както полезността, така и привлекателността на услугата (Evens и др., 2024).

Платформите за видео съдържание с кратка продължителност – като TikTok, Instagram Reels и YouTube Shorts – въвеждат нови тенденции, които променят глобалните маркетингови стратегии. Те се отличават с алгоритмично управлявани потоци от видеа, бързо създаване на съдържание и висока степен на интерактивност. Изследванията показват, че тези характеристики влияят върху ефективността на

маркетинга, тъй като алгоритмите за препоръки, моделите на създаване на съдържание и динамиката на взаимодействие определят кои видеа стават видими и как аудиторията реагира. Разбирането на тези особености е ключово за разработването на ефективни маркетингови подходи в екосистемата на кратките видеа (Huang и др., 2025). В основата на тази екосистема са системите за препоръки – усъвършенствани механизми с изкуствен интелект, които анализират поведението на потребителите, включително време на гледане, ангажираност и история на съдържанието, за да създават персонализирани потоци.

Системите за препоръки в платформите за кратко видео са усъвършенствани механизми с изкуствен интелект, които анализират поведението на потребителите, включително време на гледане, ангажираност (харесвания, коментари, споделяния) и история на консумираното съдържание. На базата на тези данни алгоритмите създават персонализирани потоци, осигурявайки висока степен на ангажираност на аудиторията чрез предоставяне на видеа, съобразени с индивидуалните предпочитания. Макар това да позволява на маркетолозите бързо да достигат до релевантни аудитории, то създава и предизвикателства, тъй като маркетинговите материали трябва да се конкурират с огромно количество потребителско съдържание. Успешните кампании използват алгоритмите чрез създаване на ангажиращо и лесно за споделяне съдържание, но ефективността зависи от внимателното съобразяване на маркетинговите стратегии с логиката на алгоритмите и поведението на потребителите (Shakurov и др., 2021).

Анализът на съдържанието и динамиката в онлайн медийното пространство все повече разчита на интеграцията на технологии за извличане на данни и методи за обработка на естествен език (NLP), обобщени в концепцията за медийна интелигентност (Милев, 2025b). Приложенията за видео съдържание могат да бъдат разгледани през четири ключови перспективи – съдържание, потребители, социално взаимодействие и услуги – които определят тяхната същност и функционалност. От гледна точка на съдържанието тези платформи позволяват на потребителите да създават и разпространяват видеа чрез качване или предаване на живо. От гледна точка на потребителите, авторите на материали споделят видеа с лична значимост, за да ангажират аудиторията. Някои от тези автори се превръщат в професионални създатели на съдържание и добиват популярност като известни личности, генериращи доходи чрез споделяне на приходи, партньорски маркетинг или дарения от зрители. Социалната перспектива акцентира върху интерактивни функции като харесване, споделяне, абониране, коментиране, чат и дарения, които насърчават създаването на виртуални общности и засилват социалната ангажираност, докато перспективата на услугите включва алгоритми за класиране, препоръки и персонализация, които подобряват доставката на съдържание и оптимизират потребителското изживяване.

Обобщено, тези четири перспективи показват, че видео платформите са сложни социотехнически екосистеми, в които съдържанието, потребителите, социалното взаимодействие и алгоритмичните услуги взаимно се влияят и оформят както функционалността, така и ангажираността на аудиторията (Bartolome и др., 2023).

Разбирането на социотехническата същност и тенденциите в развитието на видео платформите разкрива необходимостта от стабилна технологична инфраструктура, способна да поддържа мащабируемостта и функционалността на тези среди. В основата на тази инфраструктура лежат фундаментални технически процеси и стандарти, които определят как аудиовизуалното съдържание се кодира, компресира и визуализира за крайния потребител. Поради това последващият анализ се фокусира върху технологичните особености на платформите за видео споделяне, осигуряващи ефективната доставка и възпроизвеждането на медийните ресурси.

3. Технологични особености на платформите за видео споделяне

В дигиталната медийна среда видео данните трябва да бъдат ефективно представени, пренасяни и съхранявани без значителна загуба на качество. Тази ефективност се постига чрез видео кодиране – процес на компресиране на суровия материал, като същевременно се запазва визуалното качество и се осигурява съвместимост между различни устройства и платформи. За да се подпомогне този процес, са разработени различни стандарти и формати за кодиране, които балансират между степента на компресия, качеството на изображението и гъвкавостта при възпроизвеждане.

Видео кодирането (video coding) преобразува визуалните данни в компактен двоичен формат с цел ефективно съхранение и предаване. През последните десетилетия основните международни стандарти за видео кодиране са разработени съвместно от две организации: Moving Picture Experts Group (MPEG) към ISO/IEC и Video Coding Experts Group (VCEG) към ITU-T. Тези сътрудничества водят до създаването на поредица от широко използвани стандарти: H.262/MPEG-2 (1995), H.264/AVC (2003), H.265/HEVC (2013) и H.266/VVC (2020). Всички тези стандарти се основават на комбинация от беззагубни и загубни методи за компресия, както и на перцептивни модели на човешката зрителна система:

- Беззагубна компресия – намалява статистическата излишност чрез техники като ентропийно кодиране.
- Загубна компресия – включва методи като предсказване и квантизация за съществено намаляване на обема на данните.
- Перцептивна компресия – премахва информация, която не се възприема от човешкото зрение (Casella, 2025).

Докато видео кодирането определя теоретичната рамка и алгоритмите за компресия, видео енкодирането (video encoding) представлява практическата реализация на тези принципи. То определя как видео потокът се обработва, пакетира и съхранява в конкретни файлови формати.

Форматът за видео кодиране (video encoding format) се отнася до техническата структура и метода на компресия, използвани за представяне на видео файл. Той определя качеството, адаптивността и използваемостта на файла в различни платформи и устройства. Съществуват множество формати, включително AVI, DV, HDV, MOV, WMV, MPEG-1, MPEG-2, MPEG-4, MJPEG, MKV и 4K.

Суровите видео записи (raw video) заемат значително пространство за съхранение – често няколко гигабайта на минута в зависимост от резолюцията и кадровата честота. Поради това компресията е необходима за намаляване на размера на файловете с цел ефективно съхранение, разпространение или стрийминг. Методите на компресия се различават по ефективност, запазване на качеството и съвместимост, което влияе върху избора на формат според нуждите на продукцията и техническите изисквания.

Форматите за видео кодиране определят начина, по който даден филм се съхранява, достъпва и разпространява – дали чрез физически носители като VHS или VCD, чрез дигитални устройства за съхранение или чрез онлайн платформи за поточно предаване на видео. Форматът влияе също върху степента на компресия, качеството на изображението и възможностите за редактиране или възпроизвеждане. С течение на времето са разработени множество стандарти, които отговарят на развиващите се технологични изисквания и медийни приложения (Onyejelem, 2024).

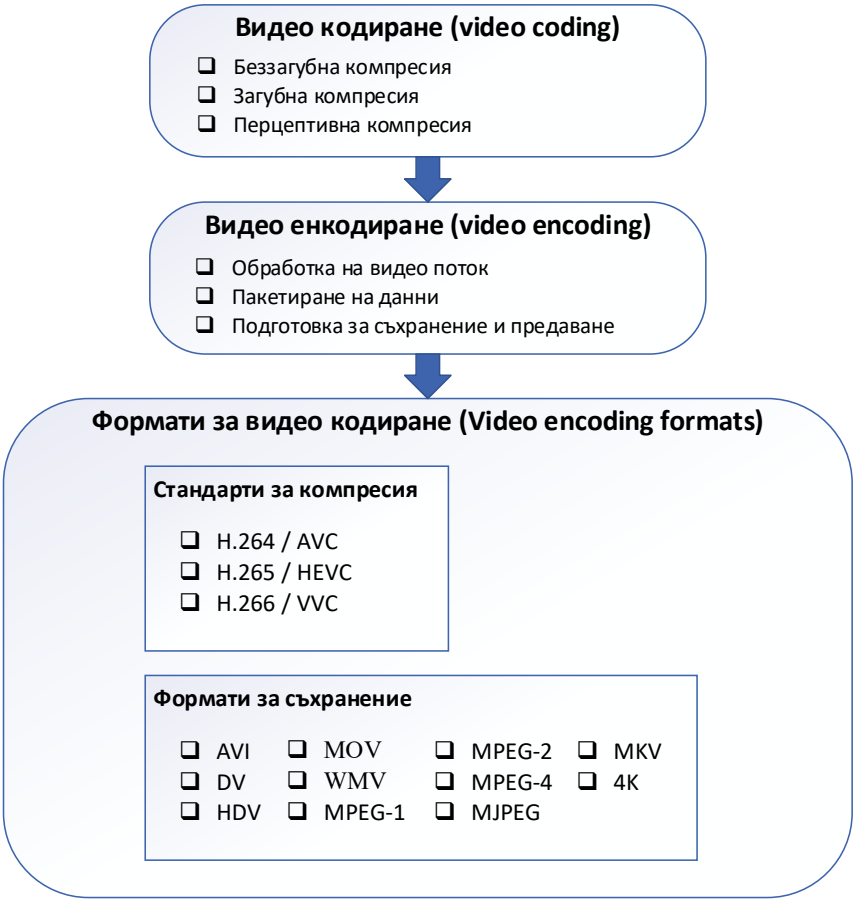
В технологичната рамка на съвременните платформи за видео споделяне, видео кодирането и форматите за кодиране играят централна роля за оптимизиране на доставката и възпроизвеждането на съдържание. Платформите използват стандартизирани алгоритми като H.264/AVC и H.265/HEVC за компресиране и транскодиране на качените видеа в различни формати и битрейтове (количеството данни, обработени за секунда). Това позволява адаптивно поточно предаване и осигурява постоянно качество при различни устройства и мрежови условия. Ефективността на избрания стандарт и формат пряко влияе върху производителността, изискванията за съхранение и цялостното потребителско изживяване.

Фигура 2 илюстрира принципите на видео кодирането и съответните формати за кодиране.

В контекста на съвременните платформи за споделяне на видео редица компоненти на потребителския интерфейс (UI) играят ключова роля за ангажираността на потребителите и достъпността на съдържанието. Компонентите на потребителския интерфейс са проектирани да осигурят ефективна доставка на медийно съдържание, интуитивна навигация и персонализирано взаимодействие.

Основните елементи на интерфейса са:

- Видео плейър – централен модул, който предоставя директен достъп и контрол върху аудиовизуалното съдържание. Въвеждането на HTML5 премахва нуждата от външни модули, като осигурява вградена под-



Фигура 2. *Принципи на видео кодирането и формати за кодиране*
Източник: *Авторово изображение*

дръжка на мултимедийни формати в браузърите и позволява гладко и адаптивно възпроизвеждане на различни устройства (Shende, 2024).

- Търсачка – позволява ефективно откриване на видеоклипове или канали чрез ключови думи, използвайки индексирани метаданни като заглавия, описания и тагове. В платформи като YouTube и Vimeo тези системи често включват персонализация според потребителските предпочитания (Reiter и др., 2021).
- Система за препоръки – персонализира достъпа до съдържание чрез предложения, базирани на предишни взаимодействия и предпочитания. Тези системи анализират поведението на потребителите и характеристиките на видео съдържанието с цел увеличаване на ангажираността, времето на гледане и задържането на аудиторията (Lubos и др., 2023).

Осигуряването на ефективно кодиране, компресия и интуитивен

потребителски интерфейс изгражда необходимата техническа основа на видео средите, но огромният обем от ежедневно качвано съдържание налага внедряването на интелигентни подходи за неговата организация. Технологичната инфраструктура трябва да бъде надградена с механизми за автоматично извличане на контекст и семантично съдържание. Водена от тази цел, следващата точка анализира принципите и моделите за автоматизирано структуриране на метаданни, с акцент върху мултимодалната интеграция на аудио, визуална и текстова информация.

4. Архитектурни и технически аспекти на видео кодирането и потребителския интерфейс

Бързото разширяване на дигиталните медии и платформите за споделяне на видео доведе до безпрецедентен ръст на мултимедийното съдържание, създавайки значителни предизвикателства по отношение на неговата организация, извличане и ефективно използване. В този контекст метаданните играят централна роля, като позволяват индексиране на съдържание, търсене и персонализирани препоръки. Въпреки това традиционните подходи, базирани на ръчно анотиране на метаданни, остават неефективни и трудни за мащабиране, особено при обработка на големи обеми видео данни. Това ограничение доведе до нарастващ интерес към техники за автоматично генериране на метаданни, които се основават на вътрешните характеристики на мултимедийното съдържание (Yong и др., 2022).

Мултимедийните данни по своята същност са хетерогенни и включват видео, аудио, текстови и визуални елементи. Сложността на такива структури изисква усъвършенствани механизми за семантично индексиране, които позволяват на системите да интерпретират съдържанието отвъд характеристиките от най-ниско ниво, като цветови, текстурни и динамични особености на видеото. Развитието на семантичната мрежа допълнително подобри представянето на знания, като позволи мултимедийните ресурси да бъдат описвани чрез структурирани, машинно четими метаданни, което повишава ефективността на търсенето и контекстуалната релевантност (Amaria и др., 2022). В този контекст автоматичното анотиране на видео се утвърди като ключова технология, която свързва описателна информация с конкретни сегменти от видео съдържанието и улеснява мащабния анализ на мултимедийно съдържание (Ambore и др., 2024).

За да се отговори на многоаспектния характер на видео данните, съвременните подходи за автоматизирано генериране на метаданни все повече разчитат на мултимодален анализ, интегриращ допълващи се перспективи, извлечени от аудио, визуална и текстова информация. Всяка модалност предоставя различни семантични сигнали, а тяхното комбинирано използване позволява по-пълно разбиране на видео съдържанието.

В този контекст анализът на аудио представлява фундаментален компонент на мултимодалното генериране на метаданни, като осигурява

достъп до лингвистична информация, вложена в звуковите пътеки на видеото. По-специално, автоматичното разпознаване на реч (ASR) позволява преобразуването на говоримия език в текстова форма, подпомагайки генерирането на транскрипции, субтитри и други езиково-базирани метаданни, които улесняват индексирането и извличането (Jurafsky и др., 2026). Системите за разпознаване на реч работят чрез идентифициране на модели в акустичните сигнали и съпоставянето им със съответните езикови единици. Тези системи се използват широко в приложения като гласово управлявани интерфейси, услуги за транскрипция и помощни технологии, при които взаимодействието на естествен език играе централна роля в комуникацията човек–компютър (Nayeem и др., 2025). От методологична гледна точка ASR разчита на техники за машинно обучение, за да моделира връзката между аудио характеристиките и езиковите структури, позволявайки извличането на смислени текстови представяния от речеви данни (Al-Fraihat и др., 2024).

Въпреки значителния напредък, анализът, базиран единствено на аудио, не може напълно да улови сложността на видео съдържанието, което налага неговото допълване с други модалности. За преодоляване на тези ограничения се използва анализът на визуално съдържание, който се фокусира върху извличането на информация директно от видео кадрите чрез техники от компютърното зрение. Тези подходи анализират вътрешни визуални характеристики като цветови разпределения, динамика на движението, структура на обекти и лицеви характеристики, което позволява идентифициране на обекти, сцени и събития във видеото.

В практиката визуалният анализ се оказва особено ефективен в области като видео архивиране и управление на съдържание, където автоматизираните системи подпомагат откриването и разпознаването на лица или обекти в големи мултимедийни колекции. Такива възможности значително намаляват ръчния труд при създаването на метаданни и подобряват ефективността на архивните процеси (Shakurov и др., 2021). Чрез прилагане на методи за семантичен анализ могат автоматично да се извличат описания на сцени, обекти и други характеристики на видеосъдържанието, които се интегрират в метаданните и подпомагат по-ефективното търсене и извличане на информация (Нап и др., 2023).

Освен това визуалните характеристики могат да улавят по-високо ниво на семантични свойства на видео съдържанието. Например представителни ключови кадри, извлечени от видео последователности, могат да отразяват наративна структура, тематични елементи или жанрови характеристики. В комбинация с допълващи сигнали като аудио модели, тези визуални дескриптори предоставят по-богата основа за класификация и интерпретация на съдържанието (Yong и др., 2022). Въпреки това, визуалният анализ сам по себе си остава недостатъчен за пълно семантично разбиране, което налага включването на текстови методи.

В тази връзка обработката на естествен език позволява извличане и

интерпретация на семантична информация от текстови данни, свързани с видео съдържание. Като основен компонент на автоматизираното генериране на метаданни, NLP се прилага върху субтитри, транскрипции и други текстови елементи, извлечени от мултимедийни източници, като трансформира неструктуриран език в структурирани представяния, подходящи за анализ.

NLP процесът обикновено включва предварителни стъпки като нормализация на текста, токенизация, стеминг (редуциране на думите до корен) и премахване на нерелевантни елементи, което позволява на системите да изолират значими лингвистични характеристики. Тези характеристики, включително ключови думи и семантични връзки, служат като основа за по-високо ниво задачи като обобщаване на съдържание, откриване на теми и класификация (Vijayalakshmi и др., 2024). Чрез използване на техники за машинно и дълбоко обучение, NLP моделите могат да идентифицират модели в големи текстови масиви и да извеждат контекстуално значение. Тази способност позволява по-прецизна категоризация на видео съдържанието и обогатява семантичната стойност на генерираните метаданни (Amaria и др., 2022). Въпреки това, подобно на аудио и визуалния анализ, NLP самостоятелно предоставя само частична перспектива, което подчертава необходимостта от интегрирани подходи.

Поради допълващия се характер на аудио, визуалната и текстовата информация, мултимодалната интеграция се утвърждава като централен подход в автоматизираното генериране на метаданни. Мултимодалните подходи комбинират техники, базирани на съдържание, които извличат характеристики директно от видео и аудио потоци, с текстово-базирани методи, използващи NLP за интерпретация на езикови данни (Ambore и др., 2024). Такива хибридни рамки позволяват на системите да улавят както структурните характеристики, така и семантичното значение на мултимедийното съдържание. Чрез корелация на информация между различните модалности — например съпоставяне на говор с визуални сцени и текстови описания — тези подходи значително подобряват точността и контекстуалната релевантност на генерираните метаданни (Yong и др., 2022).

Предимствата на мултимодалната интеграция са особено очевидни при мащабни платформи за споделяне на видео, където разнообразно и динамично съдържание трябва да бъде обработвано ефективно. Чрез използване на напредъка в машинното и дълбокото обучение, мултимодалните системи могат да автоматизират процеса на анотиране, да намалят зависимостта от ръчен труд и да подобрят ефективността на механизмите за индексиране, търсене и препоръки (Ambore и др., 2024).

На фигура 3 е представен схематичен модел на процеса на мултимодална интеграция.

В заключение, интеграцията на аудио и визуален анализ, както и обработка на естествен език, предоставя



Фигура 3. *Процес на мултимодална интеграция*
Източник: *Авторово изображение*

стабилна и мащабируема рамка за автоматизирано генериране на метаданни. Чрез комбиниране на допълващи се източници на информация, мултимодалните подходи позволяват по-дълбоко и по-цялостно разбиране на видео съдържанието, което е от съществено значение за управлението на нарастващата сложност на съвременните мултимедийни екосистеми и за подпомагане на усъвършенствани услуги за откриване и персонализация на съдържание (Amaria и др., 2022).

Заключение

В условията на динамично развиващата се дигитална среда платформите за видео споделяне се утвърждават като сложни социотехнически системи, в които ефективното управление на съдържанието е от съществено значение. Настоящото изследване подчертава централната роля на метаданните като основен инструмент за структуриране, организиране и откриване на видео ресурси, като същевременно очертава тяхното значение за подобряване на достъпността, потребителското изживяване и ефективността на информационните системи.

Анализът показва, че традиционните подходи за ръчно създаване на метаданни са ограничени по отношение на мащабируемост и ефективност, особено в контекста на непрекъснато нарастващите обеми от мултимедийно съдържание. В отговор на тези предизвикателства автоматизираното генериране на метаданни се очертава като необходима и перспективна посока, която използва вътрешните характеристики на видеото и съвременни методи от областта на изкуствения интелект.

Особено значение има мултимодалният подход, който интегрира аудио анализ, визуално разпознаване и обработка на естествен език. Чрез комбиниране на тези допълващи се източници на информация се постига по-дълбоко и по-прецизно разбиране на съдържанието, което значително подобрява качеството на генерираните метаданни и съответно ефективността на процесите по индексиране, търсене и препоръки.

Освен технологичните аспекти, изследването откроява и значението на архитектурните и функционалните характеристики на съвременните видео платформи, включително ролята на алгоритмите за персонализация, интерактивните механизми и адаптивните стрийминг технологии. Всички тези компоненти функционират в тясна връзка с метаданните, което допълнително подчертава тяхната ключова позиция в екосистемата на дигиталното видео съдържание.

В перспективен план ролята на метаданните в платформи за видео споделяне от мащаба на YouTube ще продължи да се развива под въздействието на бързо навлизащи генеративни модели с изкуствен интелект и технологии за семантичен анализ. Бъдещите тенденции сочат към преминаване от статично анотиране към динамично, автоматично генерирано описание под формата на метаданни в реално време, което отчита интерактивния контекст и потребителската ангажираност. Подобен напредък ще позволи още по-прецизно структуриране на съдържанието на ниво конкретни кадри и теми, трансформирайки платформите от пасивни хранилища в интелигентни и ориентирани към знания екосистеми.

В заключение може да се обобщи, че автоматизираното структуриране на метаданни представлява критичен фактор за устойчивото развитие на платформите за видео споделяне. Чрез прилагането на мултимодални и интелигентни подходи се създават предпоставки за по-ефективно управление на съдържанието, по-висока релевантност на предоставяната информация и

по-добра адаптация към потребителските нужди в условията на нарастваща дигитална комплексност.

Публикацията съдържа резултати от изследване, финансирано със средства от целева субсидия за НИД на УНСС по договор № НИД НИ-26/2026/Б.

Цитирани източници (References):

1. Милев, П. (2025a). Архитектурни решения за информационна система с потребителска конфигурация в университетска среда. *Икономически и социални алтернативи,* (1), 100–115.
   (Milev, P. (2025a). Arhitekturni reshenia za informatsionna sistema s potrebitelska konfiguratsia v universitetska sreda. *Ikonomicheski i sotsialni alternativi,* (1), 100–115)
2. Милев, П. (2025b). Разбиране на уеб публикациите в съвременната дигитална среда чрез приложение на концепцията за интелигентност на медиите. *Научни трудове на УНСС,* (4), 27–41.
   (Milev, P. (2025b). Razbirane na ueb publikatsiite v savremennata digitalna sreda chrez prilozhenie na kontseptsiyata za inteligentnost na mediite. *Nauchni trudove na UNSS,* (4), 27–41)
3. Al-Fraihat, D., Sharrab, Y., Alzyoud, F., Qahmash, A., Tarawneh, M. & Maaita, A. (2024). Speech recognition utilizing deep learning: A systematic review of the latest developments. *Human-Centric Computing and Information Sciences,* 14(15), 1–33. https://doi.org/10.56830/hcis.2024.14.15
4. Allen, R.B. (2020). Metadata for social science datasets. In: J.I. Lane, I. Mulvany & P. Nathan (Eds.), *Rich search and discovery for research datasets: Building the next generation of scholarly infrastructure*. Sage Publishing.
5. Amaria, S., Guidedi, K., Lazarre, W. & Kolyang (2022). A survey on multimedia ontologies for a semantic annotation of cinematographic resources for the web of data. *Acadlore Transactions on AI and Machine Learning,* 1(1), 2–10.
6. Ambore, B., Neha, B.A., Karanth, S., Sharma, G. & Saxena, S. (2024). Automation of video annotation using textual description: A comprehensive survey. *International Journal for Research Trends and Innovation,* 9(1), 513–526.
7. Bailer, W. & Höffernig, M. (2009). Metadata for creation and distribution of multi-view video content. In *Proceedings of the 10th International Workshop of the Multimedia Metadata Community on Semantic Multimedia Database Technologies (SeMuDaTe'09)* (Vol. 539, Paper 4). CEUR-WS.org. https://ceur-ws.org/Vol-539/paper_4.pdf
8. Bartolome, A. & Niu, S. (2023, April). A literature review of video-sharing platform research in HCI. In: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (pp. 1–20). ACM.
9. Casella, A. (2025). *Evolution of video codecs: A comparative analysis of HEVC, VVC, AV1, and AVM* (Doctoral dissertation, Politecnico di Torino).
10. Evens, T., Henderickx, A. & Conradie, P. (2024). Technological affordances of video streaming platforms: Why people prefer video streaming platforms over television. *European Journal of Communication,* 39(1), 3–21.

11. Formenton, D. & de Souza Gracioso, L. (2022). Metadata standards in web archiving technological resources for ensuring the digital preservation of archived websites. *Revista Digital de Biblioteconomia e Ciência da Informação,* 20, e022018.
12. Gadient, P., Nierstrasz, O. & Ghafari, M. (2021, December). Security header fields in HTTP clients. In *2021 IEEE 21st International Conference on Software Quality, Reliability and Security (QRS)* (pp. 93–101). IEEE.
13. Han, F. & Meng, Z. (2023). *A metadata generation system with semantic understanding for video retrieval in film production*. arXiv. https://arxiv.org/abs/2312.00104
14. Huang, A. & Zainol, Z. (2025). The impact of short video platform characteristics on marketing effectiveness: A bibliographic review and future research directions. *International Journal of Global Economics and Management,* 6(1).
15. Jones, S.M., Neblitt-Jones, V., Weigle, M.C., Klein, M. & Nelson, M.L. (2021, September). It's all about the cards: Sharing on social media encouraged HTML metadata growth. In *2021 ACM/IEEE Joint Conference on Digital Libraries (JCDL)* (pp. 110–119). IEEE.
16. Jurafsky, D. & Martin, J.H. (2026). *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition with language models* (3rd ed. draft). Stanford University and University of Colorado Boulder.
17. Lubos, S., Felfernig, A. & Tautschnig, M. (2023). An overview of video recommender systems: State-of-the-art and research issues. *Frontiers in Big Data,* 6, 1281614.
18. Mosha, N.F. & Ngulube, P. (2023). Metadata standard for continuous preservation, discovery, and reuse of research data in repositories by higher education institutions: A systematic review. *Information,* 14(8), 427.
19. Nayeem, M., Tabrej, M.S., Deb, K.J., Goswami, S. & Hakim, M.A. (2025). *Automatic speech recognition in the modern era: Architectures, training, and evaluation*. arXiv. https://arxiv.org/abs/2510.12827
20. Onyejelem, T.E. (2024). *Film production: Video encoding formats, theories and aesthetics*. University of Lagos Press and Bookshop Ltd.
21. Patel, B.V. & Gaharwar, R.D. (2018). Search engine optimization (SEO) using HTML meta-tags. *International Journal of Scientific Research in Science and Technology,* 4(9), 298–302.
22. Reiter, M., Cupka, O. & Miklosik, A. (2021). *Search engine optimization of video content*. ResearchGate.
23. Sathish Kumar, S. & Pushpa, P. (2019). Preparing metadata for making video file searchable on online. *Journal of Emerging Technologies and Innovative Research,* 6(3), 512–516.
24. Shakurov, M., Korobkin, D., Fomenkov, S. & Golovanchikov, A. (2021, October). Software for generation of video files metadata. *Journal of Physics: Conference Series,* 2060(1), 012029. IOP Publishing.

25. Sharma, A. & Gupta, V. (2024). An analysis of impact of metadata image, video and PDF in web applications using performance testing. *Journal of Electrical Systems,* 20(11s).
26. Shende, V.S. (2024). HTML5 in web development. In: *Recent Advancements in Science and Technology* (p. 206).
27. Vijayalakshmi, K., Fatima, M., Reddy, G.J., Reddy, M.G. & Manideep, K. (2024). A deep learning-based approach for inappropriate content and classification of YouTube videos. *International Journal for Advanced Research in Science & Technology,* 14(5), 886–894.
28. Xiang, Z., Horváth, J., Baireddy, S., Bestagini, P., Tubaro, S. & Delp, E.J. (2021). Forensic analysis of video files using metadata. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 1042–1051). IEEE.
29. Yong, S.J., Park, H.G., You, Y.H. & Moon, I.Y. (2022). Automatic generation of video metadata for the super-personalized recommendation of media. *Journal of Information and Communication Convergence Engineering,* 20(4), 288–294.

Printsipi za avtomatizirano strukturirane na metadanni v platformi za video spodelyane

Yavor Tabov

Principles for Automated Structuring of Metadata in Video Sharing Platforms

Yavor Tabov

Abstract: Rapid advancements in digital technologies and the growing importance of online video content are transforming the way information is created, distributed, and consumed. Video sharing platforms are becoming key environments for communication and knowledge exchange, placing high demands on content management and discovery. In this context, metadata plays a crucial role in structuring and optimizing video resources. This paper examines the principles of automated metadata structuring, analyzing their nature, types, and functions in a web-based environment. The study covers trends in the development of video platforms as well as technological aspects related to video coding, video compression methods, and user interface components. Particular emphasis is placed on the automated generation of metadata through multimodal analysis, which integrates audio, visual, and textual information. The conclusion highlights the importance of automated metadata structuring for the functioning and development of video sharing platforms. It also emphasizes the role of multimodal approaches in achieving more accurate description and interpretation of video content, with the aim of improving indexing, search, and personalization processes, and creating conditions for more efficient management and delivery of video content.

Key words: metadata, video sharing platforms, multimodal integration.

JEL: C88, L86.