

Влияние на интелигентността на данните върху структурирането на съдържание в уеб публикации

Пламен Милев*, Явор Табов**

Резюме: Статията представя цялостно изследване на влиянието на интелигентността на данните върху структурирането на съдържание в уеб публикации. Започвайки с изследване на различните стратегии и технологии, които са в основата на интелигентността на данните, статията подчертава трансформиращото въздействие на интелигентността на данните върху представянето и използваемостта на съдържанието в уеб публикации. Статията изследва възможностите и предизвикателствата, свързани с внедряването на интелигентност на данните в уеб публикации, включително потенциала за засилено ангажиране на потребителите, опасенията за поверителността и необходимостта от непрекъснати инвестиции в технологиите. Поради непрестанното развитие в изследваната област по проблематиката за интелигентност на данните, статията разглежда очакваното въздействие на нововъзникващите технологии. В изследването се акцентира на необходимостта от структуриране на данни в уеб публикации и процеса на

трансформиране на неструктурираните и полуструктурирани данни в структурирана форма. Подчертава се значението на структурираните данни за подобряване на потребителското изживяване, улесняване на операциите в контекста на интелигентност на данните и подобряване на практиките за управление на данните. В заключение се обобщава необходимостта от последващо проучване на стратегии и технологии за структуриране на данни, за да се осигурят възможности за приложението на интелигентност на данните в уеб публикации.

Ключови думи: интелигентност на данните, структуриране на данни, интернет, уеб концепция, онлайн съдържание.

JEL: C88, L86.

Увод

В съвременното дигитално общество концепцията за интелигентност на данните (data intelligence) представлява съществен интерес от изследователска гледна точка. Интелигентността на данните, капсулирайки широк спектър от стратегии и технологии за обработка и интерпретация на данни,

* Пламен Милев е доктор, доцент в катедра „Информационни технологии и комуникации“ на УНСС.

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

фундаментално трансформира областта на уеб публикуването. Интелигентността на данните подобрява достъпността, използваемостта и цялостната стойност на данните, представени в рамките на различните уеб базирани информационни системи, което от своя страна води до по-голяма ангажираност от страна на потребителите на съответните платформи в резултат от предоставените възможности за по-задълбочено персонализиране на съдържанието. От изследователска гледна точка интерес представляват различните аспекти на интелигентността на данните в контекста на процесите на структуриране на данни в уеб базирани информационни системи. Разглежданата в статията проблематика включва присъщите предизвикателства при приложението на интелигентност на данните в уеб публикации. Статията изследва въпроси, имащи отношение към поверителността на потребителите, разходите за поддържане на инфраструктура за интелигентност на данните и продължаващата необходимост от адаптиране към бързо променящото се технологично развитие. Изследването включва и проучване на трансформационния потенциал на нововъзникващите технологии върху бъдещето на приложението на интелигентност на данните в уеб публикации. Статията подчертава значението на структурирането на данни в уеб базирани информационни системи. В хода на изследването се подчертава как представянето на неструктурирани и полуструктурирани данни в структурирана форма улеснява по-ефективното управление на данните, подобрява потребителските възможности и позволява приложението

на различни подходи за интелигентност на данните върху съответното съдържание. В статията подробно се представя процесът на трансформиране на данни, като се очертават различните използвани техники и технологии.

Целта на настоящото изследване е да подчертае значението на възприемането на балансиран подход към приложението на интелигентност на данните, с което да се максимизират ползите от тази концепция, като същевременно се намалят съответните потенциални рискове. В резултат от разглежданата изследователска проблематика могат да се дефинират препоръки и предложения за бъдещи изследвания, което да доведе до етичното и ефективно приложение на концепцията за интелигентност на данните в уеб публикации. Статията има за цел и да определи ролята на интелигентността на данните в контекста на необходимостта от представянето на съдържанието на уеб публикации в структурирана форма.

1. Интелигентност на данните като концепция

С развитието си интелигентността на данните се очертава като критична сила, открояваща необходимостта от структурирането на данни в уеб публикации. В дигитална ера, характеризираща се с нарастваща роля на данните, от съществено значение е да се проучи как интелигентността на данните трансформира областта на уеб публикуването, оптимизирайки използваемостта и достъпа до данни. Интелигентността на данните включва стратегии, инструменти и технологии за обработка и интерпретиране на данни за улесняване на

Вземането на решения (Chen и гр., 2012). Уеб публикациите, от друга страна, представляват онлайн съдържание, публикувано в различни видове уеб базирани информационни системи под формата на постове в блогове, новинарски статии, електронни книги, научни статии и гр. Настоящото изследване се фокусира върху изследването на това как интелигентността на данните влияе върху структурирането на съдържанието в уеб публикациите в съответните уеб базирани информационни системи.

Интелигентността на данните е усъвършенствана форма на бизнес интелигентност, която обхваща множество аспекти като обучение на данни, анализ на данни и визуализация на данни. Това включва използването на управлявани от изкуствен интелект (AI) технологии за обработка, анализиране, интерпретиране и визуализиране на данни, което позволява на организациите да извличат прозрения, които могат да подкрепят стратегически решения (Fayyad и гр., 1996). Използвайки интелигентността на данните, платформите за публикуване на уеб съдържание могат да оптимизират структурирането и представянето на своите данни, за да подобрят потребителските изживяване и ангажираност.

В рамките на настоящото изследване ще разглеждаме интелигентността на данните като концепция и по-специално широка концепция, която обхваща няколко аспекта на обработка, управление, анализиране и използване на данни (фигура 1).

Интелигентността на данните може да се представи като концепция, която се отнася до практиките и технологиите, които организациите от публичния и частния сектор използват, за да разберат, анализират и използват своите данни ефективно. Това може да включва широк набор от техники и подходи – от традиционен статистически анализ до по-напреднали методи за машинно обучение и изкуствен интелект. Целта на интелигентността на данните е да превърне необработените данни във фактически прозрения, които могат да подпомогнат вземането на управленски решения и стратегическото планиране в съответните организации. За да се постигне това, интелигентността на данните включва няколко взаимосвързани дисциплини като управление на данни, качество на данните, интегриране на данни, анализ на данни, визуализация на данни и гр. В контекста на настоящото изследване е важно и да се отбележи, че интелигентността на данните не се



Фигура 1. Концепция за интелигентност на данните

отнася само до техническите аспекти на обработката и анализа на данните. Интелигентността на данните също така включва съображения за поверителност на данните, сигурност и спазване на съответните правила, норми и разпоредби. В този смисъл следва да се гарантира, че данните се използват не само ефективно, но и отговорно, а също така и етично. В допълнение на изследователската гледна точка, че интелигентността на данните следва да се разглежда като многостранна концепция, интелигентността на данните може да се представи и като подход за използване на данни, който съчетава технология, методология и стратегия за извличане на ценни прозрения за данни (data insights) и вземане на информирани управленски решения.

Един от значимите начини, по които интелигентността на данните оказва влияние върху структурирането на данните в уеб публикации, е чрез подобряване на персонализацията. Персонализирането, улеснено от интелигентността на данните, включва специфициране на уеб съдържание въз основа на потребителско поведение, предпочитания и контекст (Mico, 2019). Персонализираните структури от данни отговарят на нуждите на индивидуалните потребители, като подобряват потребителското изживяване и повишават нивата на ангажираност. Освен това интелигентността на данните има отношение към семантичното структуриране на данните – съществен компонент в интерпретацията на данни. Чрез семантично структуриране данните се организират във формати, които позволяват лесен анализ, подобрявайки тяхната четимост

и интерпретируемост (Hendler, 2001). Семантичните технологии, подкрепени от интелигентни данни, помагат при създаването на контекстно подходящо съдържание, повишавайки използваемостта на уеб публикациите. Интелигентността на данните също влияе върху динамичното структуриране на данните в уеб публикации. Динамичното структуриране включва организиране на данни по начини, които отговарят на промените в реално време. В този смисъл уеб публикацииите могат да използват интелигентност на данните, за да структурират своите данни въз основа на потребителски взаимодействия в реално време или чувствителни към времето събития (Fosso Wamba и др., 2015).

Някои примери подчертават влияние на интелигентността на данните върху структурирането на съдържание в уеб публикации. Компании като Google и Facebook използват интелигентни данни, за да структурират своето съдържание. Google използва своя алгоритъм RankBrain, за да персонализира резултатите от търсенето въз основа на поведението на потребителите (Selvaraj, 2019). По подобен начин Facebook използва интелигентност на данните, за да персонализира емисиите с новини, повишавайки ангажираността на потребителите (Zhang и др., 2016). В този контекст The New York Times, водеща онлайн новинарска медия, използва интелигентност на данните, за да структурира своето съдържание. Компанията използва техники, управлявани от изкуствен интелект, за да анализира поведението на читателите, като по този начин тя структурира своите данни, за да се погрижи за индивидуалните предпочитания на

читателя (Molla, 2018). Интелигентността на данните може да се разглежда и в контекста на дигиталната трансформация, която от някои автори се представя като интеграция на новите дигитални технологии във всички сфери на бизнеса (Лазарова, 2019). В този контекст има и други изследвания, които акцентират на значимостта на структурите от данни за бизнеса и организациите от публичния сектор, включително в сферата на висшето образование (Кирилов, 2019). Много често тези структури от данни прерастват в големи данни, за които авторите в предметната област причисляват осигуряването на анализи, с помощта на които се предоставят възможности за справяне с много проблеми и предизвикателства (Боянов, 2022). В сферата на висшето образование съхранението на информация в структурирани бази от данни е един от много изследваните и актуални въпроси в контекста на управлението на бизнес процеси (Белев, 2017).

В контекста на настоящото изследване следва да се отбележи, че интелигентността на данните може значително да влияе върху структурирането на съдържание в уеб публикациите чрез улесняване на персонализирането, семантичното структуриране и динамичното структуриране, за да се постигне подобряване на изживяването на потребителите, повишаване на ангажираността и оптимизиране на уместността на уеб съдържанието. Тъй като данните продължават да се превръщат в доминираща сила в цифровата ера, влиянието на интелигентните данни върху уеб публикациите вероятно ще нарасне, оформяйки

бъдещето на структурирането на данни в онлайн пространството.

2. Значение на представянето на данните в структурирана форма

Данните варират от структурирани, полуструктурирани до неструктурирани формати, всеки с различни характеристики и приложения. Фокусът на настоящото изследване е поставен върху разбирането на връзката между неструктурираните и полуструктурираните данни в контекста на потенциалното приложение на концепцията за интелигентност на данните. Неструктурираните данни, според изследване на Најај (2018), обхващат всички данни, които не се вписват в таблица на база от данни, като имейли, видеоклипове, публикации в социални медии и др. От друга страна, полуструктурираните данни, както е описано от Chang (2015), съдържат известно ниво на организация, но не съответстват на ясната структура на моделите на данни, свързани с релационни бази от данни. Както неструктурираните, така и полуструктурираните данни са присъщи на съвременните екосистеми от данни и често съществуват в тандем. Корелацията произтича от тяхното несъответствие с традиционните модели на данни, нуждата им от уникално съхранение, техники за управление и споделяната им приложимост в напредналите анализи (Agrawal и др., 2012). Разбирането на сложните връзки между неструктурираните и полуструктурираните данни изисква поглед от по-отблизо на техните характеристики и начина, по който се третират в различни аналитични сценарии. Неструктурираните данни често

са обемисти, като източници като социални медии и имейли генерират огромно количество данни ежедневно (IBM, 2021). Този тип данни, въпреки своята сложност, съдържа богати, подробни и специфични за контекста прозрения, които често е невъзможно да се получат от структурирани данни. Предизвикателството при използването на неструктурираните данни се крие в присъщата им сложност и липса на структура, което затруднява анализирането с помощта на традиционни инструменти за анализ на данни (Rajan, 2018). Полуструктурираните данни, макар и не толкова сложни като неструктурираните данни, също представляват съществено предизвикателство. Тяхната организация не е толкова ясна, колкото тази на структурираните данни, но все пак има някакво подобие на структура, често под формата на етикети и маркери, които дефинират определени елементи от данни (Chang, 2015). Пресечната точка между неструктурираните и полуструктурираните данни е в тяхното несъответствие с традиционните структури на данни и специфичните подходи, необходими за тяхното обработване и последващ анализ. И двата типа данни могат да се възползват от техники като машинно обучение и обработка на естествен език, подчертавайки значението на интелигентността на данните при обработката на сложни типове данни (Raghupathi и др., 2014).

Появата на интелигентността на данните предложи нов подход към структурирането на данни, трансформирайки начина, по който потребителите взаимодействат с уеб публикациите. Въпреки това огромният потенциал на интелигентността на данни остава

недостатъчно проучен в много сектори. Информационните системи за публикуване на уеб съдържание могат да увеличат максимално използването на интелигентни данни чрез приоритизиране на ориентирани към потребителя подходи за структурирането на данните. Това включва интегриране на механизми за обратна връзка за по-добро разбиране на потребителските предпочитания и непрекъснато усъвършенстване на техните структури от данни въз основа на тази обратна връзка. Освен това създателите на уеб базирани информационни системи следва да инвестират в най-новите технологии и инструменти за интелигентност на данните, за да се възползват от предимствата, които съответните технологии и инструменти предлагат. Изкуственият интелект, машинното обучение и предсказуемият анализ могат да бъдат използвани за оптимизиране на структурирането на данните, предлагайки по-персонализирано, ангажиращо и удобно за потребителя изживяване. Тъй като сигурността на данните и поверителността са от първостепенно значение в дигиталната ера, платформите за публикуване на уеб съдържание следва да гарантират, че тяхното използване на интелигентност на данните е в съответствие със съответните закони и етични стандарти. Това включва осигуряване на анонимност на данните, получаване на съгласие на потребителите и прозрачно съобщаване на практиките за използване на данни на потребителите.

Интегрирането на интелигентността на данните в уеб публикации носи множество последици за различни заинтересовани страни. За платформите за

създаване на уеб съдържание това представлява възможност за използване на данни по начини, които подобряват използваемостта, ангажираността и уместността на тяхното съдържание. За потребителите интеграцията на тази концепция представя по-персонализирано и интерактивно изживяване, измествайки парадигмата от пасивно потребление на съдържание към активно ангажиране. Тази интеграция обаче носи и потенциални предизвикателства, особено по отношение на поверителността и сигурността. Тъй като платформите за създаване на уеб съдържание използват интелигентни данни, те следва да се справят с въпроси като поверителност на данните, киберсигурност и етични съображения около използването на данни (Danezis и др., 2010). Като се има предвид бързият напредък в технологиите за интелигентност на данните, са необходими повече изследвания, за да се изследват по-нататък техните последици за уеб публикуването. Областите, които заслужават проучване, включват въздействието на специфични технологии за интелигентност на данните върху ангажираността на потребителите, етичните измерения на използването на интелигентността на данни в уеб публикуването и съответно как концепцията за интелигентност на данните може да бъде използвана за подобряване на приобщаването и достъпността в уеб публикациите. Освен това, тъй като тази предметна област продължава да се развива, ще бъде изключително важно да се изследват нови подходи към структурирането на данни, които използват нововъзникващи технологии. Това ще помогне да се предвидят бъдещите тенденции и възможности при

използването на интелигентни данни в сферата на уеб публикуването.

Въпреки че влиянието на интелигентността на данни върху структурирането на данните в уеб публикациите е като цяло положително, то не идва без предизвикателства. Ключова възможност е потенциалът за по-добро ангажиране на потребителите, тъй като съдържанието може да бъде съобразено с предпочитанията на отделните потребители (Hussain и др., 2018). Освен това интелигентността на данните предоставя възможност за платформите за създаване на уеб съдържание да рационализират своето съдържание, предоставяйки по-фокусирана и подходяща информация на своите потребители. Използването на интелигентни данни въвежда нови сложности. Опасения за поверителността на потребителите могат да възникнат поради значителното количество лични данни, необходими за улесняване на персонализирани уеб изживявания (Maier-Schönberger и др., 2013). Внедряването на инфраструктура за интелигентност на данните може да бъде скъпо, а в същото време бързият темп на технологична промяна налага непрекъснати инвестиции в инфраструктура.

Тъй като концепцията за интелигентност на данните продължава да се развива, нейното влияние върху структурирането на данните в уеб публикациите вероятно ще се увеличи. Нововъзникващите технологии, като машинно обучение и изкуствен интелект, ще играят ключова роля в тази проблематика. Тези технологии могат да улеснят по-усъвършенствано и интуитивно структуриране на данни, като допълнително подобряват изживяването на потребителите

(Russell и др., 2016). В бъдеще може да станем свидетели на нарастващо сближаване между интелигентността на данните и други технологични тенденции. В този смисъл интегрирането на блокчейн технологията може да осигури повишена сигурност и доверие в процесите на обработка на данни. Освен това растежът на квантовите изчисления може да даде възможност за по-сложен анализ на данни и стратегии за структуриране.

3. Приложение на интелигентност на данните в уеб публикации

От критична гледна точка интелигентността на данните в уеб публикации може да се определи и като нож с две остриета. От една страна, тази концепция позволява по-динамично, персонализирано потребителско изживяване, което е от полза за ангажираността на потребителите и цялостното им удовлетворение. От друга страна обаче, това отваря вратата за потенциална злоупотреба или неправилно боравене с потребителски данни, създавайки редица етични и правни опасения. Като е формулирано от Kitchin (2014), инструментите за интелигентност на данните могат неволно да се превърнат в инструменти за наблюдение, използващи лична информация за печалба. Ако не се третираят внимателно, такива практики могат да доведат до загуба на доверие от потребителите, опетнявайки репутацията на съответната уеб платформа. Освен това приложението на изкуствен интелект и алгоритми за машинно обучение може да доведе до липса на прозрачност в начина, по който данните се обработват и представят. Този феномен

на „черна кутия“ може да създаде несигурност и потенциално да допринесе за разпространението на дезинформация, особено в контексти като новинарски публикации (Pasquale, 2015).

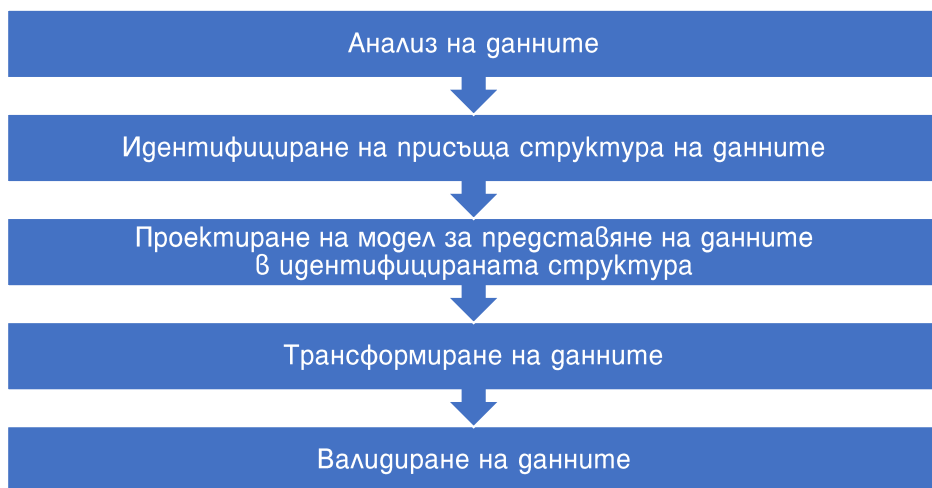
Като се имат предвид тези сложности, от решаващо значение е да се разработят политики и насоки, които да управляват използването на интелигентност на данните в уеб публикации. Тези политики трябва да имат за цел да защитават поверителността на потребителите, да насърчават прозрачността и да гарантират етични практики за данни. Възможните мерки могат да включват анонимизиране на потребителските данни, прилагане на механизми за съгласие и осигуряване на прозрачност при обработката и използването на данни. Освен това трябва да се установят индустриални стандарти и най-добри практики, които да насочват платформите за публикуване на уеб съдържание при използването на интелигентност на данните. Регулаторните органи могат да играят ключова роля в разработването и прилагането на такива стандарти, като помагат да се осигури баланс между иновациите и защитата на потребителите.

Структурирането на данните в уеб публикации е критично поради множество причини. Първо, структурираните данни подобряват използваемостта и достъпността на информацията за потребителите. В този смисъл в новинарска статия, структурирането на данни може да включва категоризиране на информация под заглавия, като по този начин улеснява потребителите да намерят и усвоят информацията, от която се нуждаят (Shneiderman и др., 1997). Освен това структурираните данни се

обработват и анализират по-лесно от инструменти за интелигентност на данните. Това е така, защото тези инструменти разчитат на подреденото подреждане на данните, за да изпълняват функциите си ефективно. В този контекст инструмент за прогнозен анализ може да изисква данните да бъдат структурирани в определен формат, за да генерират точни прогнози (Provost и др., 2013). Структурираните данни позволяват по-ефективно съхранение и извличане на данни, което може да допринесе за подобрена производителност на уеб публикациите. Това е особено важно предвид нарастващия обем данни, генерирани в цифровата ера (Zikopoulos и др., 2011). Процесът на трансформиране на неструктурирани и полуструктурирани данни в структурирана форма включва няколко стъпки (фигура 2).

Първоначалната стъпка е да се анализират данните, за да се разбере присъщата им структура. Неструктурираните текстови данни могат да съдържат

присъщи категории или теми, които могат да бъдат идентифицирани чрез техники като извличане на текст или обработка на естествен език (Feldman и др., 2007). След като присъщата структура бъде идентифицирана, следващата стъпка е да се проектира модел за улавяне на тази структура. Това може да включва създаване на таблици, категории или други видове структури, които могат да репрезентират данните. За тази цел могат да се използват инструменти като XML или JSON (Abiteboul и др., 1999). Следвайки модела на схемата, неструктурираните или полуструктурираните данни след това се трансформират в структуриран формат. Този процес, често наричан почистване на данни, може да включва набор от операции, от прости задачи като премахване на неподходящи данни до по-сложни като хармонизиране на различни източници на данни (Kandel и др., 2011). Накрая трансформираните данни се валидират, за да се гарантира, че са правилно структурирани и готови



Фигура 2. Процес на трансформиране на неструктурирани и полуструктурирани данни в структурирана форма

за използване от инструменти за интелигентност на данните. Тази стъпка е от решаващо значение за поддържане на точността и надеждността на данните в структурирана форма.

За концепцията интелигентност на данните наличието на структурирани данни е от първостепенно значение. Тъй като интелигентността на данните основно се фокусира около интерпретирането и извличането на значими прозрения от данните, форматът на данните значително влияе върху ефективността и точността на резултатите. Структурираните данни позволяват на алгоритмите да интерпретират данните по-ефективно, намалявайки времето за обработка и повишавайки точността (Нап и др., 2011). Структурираните данни позволяват и по-усъвършенствани приложения на интелигентност на данните, като машинно обучение и предсказуем анализ. Алгоритмите за машинно обучение изискват структурирани входни данни за идентифициране на модели и правене на прогнози (Alpaydin, 2020). В този контекст преобразуването на неструктурирани и полуструктурирани данни в структурирана форма значително разширява възможностите за приложение на инструментите за интелигентност на данните.

Заключение

Интелигентността на данни оказва значително влияние върху необходимостта от представянето на съдържанието в уеб публикациите в структурирана форма. Докато навлизаме в ера, в която данните все повече оформят цифровите изживявания, способността за използване на интелигентни данни се превръща

в незаменим актив за платформите за публикуване на уеб съдържание. Като позволява по-персонализирано, динамично и ориентирано към потребителя съдържание, интелигентността на данните е концепция, която е настроена да трансформира областта на публикуването в мрежата, допринасяйки за бъдеще, в което данните не просто се консумират, но активно се преживяват. Интелигентността на данните играе и трансформираща роля в структурирането на данните в уеб публикациите, като позволява по-персонализиран, динамичен и ориентиран към потребителя подход към достъпа до съдържание. Тъй като предметната област продължава да се развива, за платформите за създаване на уеб съдържание ще бъде изключително важно да бъдат в крак с най-новите постижения и да изследват новаторски начини за използване на интелигентността на данните в своите процеси по управление и обработка на данните. Въпреки че предоставя значителни възможности за подобрена ангажираност и изживяване на потребителите, тази концепция също така поставя предизвикателства, свързани с поверителността, прозрачността и етиката. Поради това е наложително да се възприеме обмислен и балансиран подход, като се използват предимствата на интелигентността на данните, като същевременно се вземат мерки във връзка с потенциалните опасности.

Структурирането на данни в уеб публикации е жизненоважно за подобряване на потребителското изживяване, улесняване на операциите за интелигентност на данните и подобряване на практиките за управление на данни. Въпреки че трансформирането на неструктурирани

и полуструктурирани данни в структурирана форма може да бъде сложно, това е решаваща стъпка в приложението на концепцията за интелигентност на данните в уеб публикациите. Тъй като тази проблематика продължава да се развива,

ще бъде важно да се изследват нови стратегии и технологии за структуриране на данни, гарантиращи успешното прилагане на концепцията за интелигентност на данните в уеб публикации.

Цитирани източници (References):

1. Белев, И. (2017). Софтуерни системи за управление на бизнес процеси и възможностите за внедряването им във висшите училища в България. *Икономически и социални алтернативи*, бр. 3, с. 96-105.
(Belev, I. (2017). Softuerni sistemi za upravlenie na biznes protsesi i vazmozhnostite za vnedryavaneto im vav visshite uchilishta v Bulgaria. *Ikonomicheski i sotsialni alternativi*, br. 3, s. 96-105)
2. Боянов, Л. (2018). Възможности и опасности от внедряването на решения в Интернет на обектите. *Икономически и социални алтернативи*, бр. 2, с. 114-120.
(Boyanov, L. (2018). Vazmozhnosti i opasnosti ot vnedryavaneto na reshenia v Internet na obektite. *Ikonomicheski i sotsialni alternativi*, br. 2, s. 114-120)
3. Кирилов, Р. (2019). Изграждане на единен регистър на работодателите на студентистажанти от УНСС. *Икономически и социални алтернативи*, бр. 4, с. 11-23.
(Kirilov, R. (2019). Izgrazhdane na edinen registar na rabotodatelite na studenti-stazhanti ot UNSS. *Ikonomicheski i sotsialni alternativi*, br. 4, s. 11-23)
4. Лазарова, В. (2019). Дигитализация и дигитална трансформация в счетоводството. *Икономически и социални алтернативи*, бр. 2, с. 97-106.
(Lazarova, V. (2019). Digitalizatsia i digitalna transformatsia v schetovodstvoto. *Ikonomicheski i sotsialni alternativi*, br. 2, s. 97-106)
5. Abiteboul, S., Buneman, P., & Suci, D. (1999). Data on the web: from relations to semistructured data and XML. Morgan Kaufmann.
6. Agrawal, D., Das, S., & El Abbadi, A. (2012). Big Data and Cloud Computing: Current State and Future Opportunities. EDBT/ICDT 2012 Joint Conference, Berlin, Germany.
7. Alpaydin, E. (2020). Introduction to machine learning. MIT Press.
8. Chang, F. (2015). Bigtable: A Distributed Storage System for Structured Data. ACM Transactions on Computer Systems.
9. Chen, H., Chiang, R.H., & Storey, V.C. (2012). Business intelligence and analytics: from big data to big impact. *MIS Quarterly*, 36(4), 1165-1188.
10. Danezis, G., & Gürses, S. (2010). A critical review of 10 years of privacy technology. Proceedings of surveillance cultures: a global surveillance society, 1-16.
11. Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). From data mining to knowledge discovery in databases. *AI Magazine*, 17(3), 37-54.
12. Feldman, R., & Sanger, J. (2007). The text mining handbook: advanced approaches in analyzing unstructured data. Cambridge University Press.

13. Fosso Wamba, S., Akter, S., Edwards, A., Chopin, G., & Gnanzou, D. (2015). How 'big data' can make big impact: Findings from a systematic review and a longitudinal case study. *International Journal of Production Economics*, 165, 234-246.
14. Hajaj, C. (2018). Understanding Unstructured Data: The wild wild west of data. Medium.
15. Han, J., Kamber, M., & Pei, J. (2011). Data mining: concepts and techniques. Elsevier.
16. Hendler, J. (2001). Agents and the Semantic Web. *IEEE Intelligent Systems*, 16(2), 30-37.
17. Hussain, A., Park, J.H., & Jung, K. (2018). A study on the user behavior in the age of personalization: Focusing on news consumption behavior. *Information Development*, 34(5), 475-488.
18. IBM (2021). What is Unstructured Data? IBM Cloud Education.
19. Kandel, S., Paepcke, A., Hellerstein, J., & Heer, J. (2011, October). Wrangler: Interactive visual specification of data transformation scripts. In Proceedings of the SIGCHI conference on human factors in computing systems (pp. 3363-3372).
20. Kitchin, R. (2014). The data revolution: Big data, open data, data infrastructures and their consequences. Sage.
21. Mayer-Schönberger, V., & Cukier, K. (2013). Big data: A revolution that will transform how we live, work, and think. Houghton Mifflin Harcourt.
22. Micó, J.L. (2019). Personalization in interactive multimedia through big data. *El Profesional de la Informacion*, 28(1), e280101.
23. Molla, R. (2018). How The New York Times uses predictive analytics to identify potential subscribers. Retrieved from <https://www.vox.com/2018/2/9/16972924/new-york-times-predictive-analytics-subscribers>
24. Pasquale, F. (2015). The black box society: The secret algorithms that control money and information. Harvard University Press.
25. Provost, F., & Fawcett, T. (2013). Data science for business: What you need to know about data mining and data-analytic thinking. O'Reilly Media, Inc.
26. Raghupathi, W., & Raghupathi, V. (2014). Big data analytics in healthcare: promise and potential. *Health Information Science and Systems*, 2, 3.
27. Rajan, T. (2018). The Challenges of Analyzing Unstructured Data. Towards Data Science.
28. Russell, S., & Norvig, P. (2016). Artificial intelligence: A modern approach. Malaysia; Pearson Education Limited.
29. Selvaraj, R. (2019). Google RankBrain: The Definitive Guide. Retrieved from <https://backlinko.com/google-rankbrain-seo>
30. Shneiderman, B., Byrd, D., & Croft, W.B. (1997). Clarifying search: a user-interface framework for text searches.
31. Zhang, M., Ackerman, M., Adamic, L., Nam, K.K., & Cheng, J. (2016). Quizz: targeted crowdsourcing with a billion (potential) users. In Proceedings of the 25th International Conference on World Wide Web (pp. 143-154).

Vliyanie na inteligentnostta na dannite varhu strukturirano na sadarzhaniye v ueb publikatsii

Plamen Milev, Yavor Tabov

Influence of Data Intelligence on the Structuring of Content in Web Publications

Plamen Milev, Yavor Tabov

Abstract: The article provides a comprehensive examination of the influence of data intelligence on the structuring of content in web publications. Beginning with an exploration of the different strategies and technologies underpinning data intelligence, the article highlights the transformative impact of data intelligence on the presentation and usability of information in web publications. The article goes on to explore the opportunities and challenges associated with implementing data intelligence in web publications, including the potential for enhanced user engagement, privacy concerns, and the need for continuous investments in technology. As the field of data intelligence continues to evolve, the article discusses the anticipated impact of emerging technologies. Furthermore, the article delves into the need for structuring data in web publications and the process of transforming unstructured and semi-structured data into structured data. It underscores the significance of structured data in enhancing user experience, facilitating data intelligence operations, and improving data management practices. The article concludes by underlining the need for further exploration of strategies and technologies for data structuring to ensure the application of data intelligence in web publications.

Key words: data intelligence, data structuring, internet, web concept, online content.

JEL: C8, L86.