

## **ФОРМАЛИЗАЦИЯ УРОВНЯ ЗРЕЛОСТИ НОВЫХ ТЕХНОЛОГИЙ В КРЕДИТНО-ФИНАНСОВОМ СЕКТОРЕ В УСЛОВИЯХ ИСПОЛЬЗОВАНИЯ СТРУКТУРИРОВАННЫХ И НЕСТРУКТУРИРОВАННЫХ КРОСС-КАНАЛЬНЫХ ИНФОРМАЦИОННЫХ РЕСУРСОВ**

© Л.Р. Магомаева

*ГГНТУ имени акад. М.Д. Миллионщикова, Грозный, Россия*

Широкое развитие индустрии больших данных в середине 2000-х годов было предопределено масштабными исследованиями в области нейтронной физики и новых методов развития искусственного интеллекта. В работе исследуется основная научная гипотеза тех времен, которая заключалась в том, что уже используемые в точных науках инструменты и методы могут быть адаптированы для преобразования структурированных данных в бизнес-аналитику и полезную информацию. Автором определяется, что в отношении неструктурированных данных используемые методы просто не работают, в результате чего возникла необходимость применения новых, ранее не изученных подходов. Также, делается вывод, что исследования последних лет, безусловно, доказывают, что сегодня идет активный поиск новых инструментов для расшифровки и адаптации неструктурированных данных.

**Ключевые слова.** Финансово-кредитный сектор, структура управления, большие данные, кросс-канальные ресурсы

Термин «структурированные кросс-канальные данные» обычно относят к данным, которые имеют определенный объем и структуру информации, зачастую она включает в себя числа, даты и группы слов и чисел, объединенных в строки. На долю такой информации приходится не больше 15-20 процентов всех данных, использование которых носит регулярный характер. По мнению некоторых авторов [8], эволюция технологий обеспечивает регулярный рост структурированных кросс-канальных данных, которые создаются в режиме реального времени.

Такие источники данных делятся на две категории: генерируемые компьютером или машиной – машинные данные, которые создаются без вмешательства человека и данные с участием человеческого труда, создаваемые в условиях взаимодействия с компьютером и информационными системами.

Примеры машинных кросс-канальных данных, образующих структурированную информацию, могут включать:

- Данные датчиков. Информация, полученная из радиочастот, интеллектуальные счетчики, медицинские устройства и данные глобальной системы позиционирования. Развитие таких информационных данных предопределено необходимостью ростом управления цепочками поставок и управления запасами.

- Данные веб-журнала. Функционирование серверов, приложений, сетей предопределяет автоматический захват определенного массива данных об их деятельности. Безусловно, это огромные объемы данных, которые могут быть полезны, например, для заключения соглашений по обслуживанию или для прогнозирования нарушений безопасности.

- Данные о точке продажи. Например, когда кассир просматривает штрих-код любого продукта, который вы покупаете, все эти данные, связанные с продуктом, генерируются.

- Финансовые данные. Большинство финансовых систем в настоящее время представляют собой полноценный программный продукт, управляемый на основе определенных правил, которые автоматизируют процессы. Ярким примером такого продукта вступает система по торговле акциями и производными финансовыми инструментами, которая содержит огромное количество структурированных данных. Некоторые из этих данных генерируются машиной, другая часть – человеком.

Примеры структурированных кросс-канальные данные, генерируемых человеком, могут включать следующую информацию:

- Входные данные. Это часть данных, которые трансформируются в структурированную информацию о человеке, например, имя, возраст, доход, ответы на вопросы, не относящиеся к свободной форме, и так далее. Эти данные могут быть полезны для понимания базового поведения клиентов в кредитной организации.

- Данные полученные по потоку уже существующих данных, генерируемых при использовании ссылки на веб-сайте. Эти данные могут быть проанализированы для определения поведения клиентов и моделей покупок.

- Данные, перемещаемые пользователями через игровой профиль.

Исследования, проведенные группой ученых во главе с Чару [4], позволили сделать обоснованный вывод о том, что большая часть структурированных кросс-канальных данных может быть получена в режиме реального времени, который полезен для понимания исходной информационной модели, подлежащей последующему анализу и прогнозированию результатов.

В отличие от структурированных данных, неструктурированные, кросс-канальные данные представляют собой огромный пласт разрозненных активов, не упорядоченных в единый информационный массив, однако именно эти данные составляют сегодня около 80-90% всей информации мировой масс медиа индустрии [6].

Некоторые эксперты и специалисты [12] в области финансовой информации убеждены, что неструктурированные кросс-канальные данные являются «черным ящиком» для единого информационного пространства, который нуждается в очистке и глубоком анализе, однако даже критики неструктурированных данных убеждены в необходимости использования передовых методов для расшифровки этой информации в целях получения новых продуктов или услуг для потребителей.

Большинство теоретиков и практиков [2], исследовавших проблемы получения кросс-канальных информационных данных, объективно считают, что большинство людей в социуме находятся в облаке неструктурированной информации, которая может быть сгенерирована или уже генерируется человеком. Яркими примерами сгенерированной информации выступают следующие источники:

- Спутниковые изображения. Сюда входят данные о погоде или данные, которые фиксируются в снимках спутникового наблюдения.

- Научные данные. Сюда относятся сейсмические изображения, атмосферные данные и физика высоких энергий.

- Фотографии и видео. Включают в себя безопасность, наблюдение и трафик.

- Данные радаров. Сюда относятся автомобильные, метеорологические и океанографические сейсмические профили.

В следующем списке приведем лишь несколько примеров неструктурированных кросс-канальных данных, генерируемых при участии человека:

- Текст, внутренняя информация о деятельности компании. В действительности корпоративная информация фактически содержит самый большой объем текстовой информации в мире.

- Данные в социальных сетях: эти данные генерируются на платформах социальных сетей, таких как YouTube, Facebook, Twitter, LinkedIn и Flickr.

- Мобильные данные. Сюда входят такие данные как текстовые сообщения и информация о местоположении.

- Содержимое веб-сайта. Генерирование этой информации происходит с любого сайта, предоставляющего неструктурированный контент, например, YouTube, Flickr или Instagram.

Безусловно, неструктурированные кросс-канальные данные являются самой большой частью информационного поля и варианты их использования в практической деятельности постоянно растут. Текстовая аналитика может использоваться для анализа неструктурированного текста и для извлечения соответствующих данных и преобразования

этих данных в структурированную информацию, которая может использоваться различными способами.

Сегодня в финансовой индустрии наиболее популярный источник для пополнения информации о клиентах является аналитика в социальных сетях, сбор и обработка которой в дальнейшем позволяют выстроить клиентские связи. Кроме того, достаточно активно используется анализ неструктурированные данные полученных из электронных сообщений, письменных комментариев, опросов и других информационных контентов для понимания модели поведения клиентов.

Некоторые авторы отмечают [3], что основным сигналом для поиска неструктурированных кросс-канальных данных является только один из многочисленных источников, информацию из которого необходимо дифференцировать на различные компоненты. В научной практике это явление принято называть «извлечение сущности». Таким образом, понимание объекта для поиска кросс-канальной информации – это первый шаг к созданию неструктурированных данных.

По утверждению Хайдмана [7], социальные сети как глобальный источник неструктурированной кросс-канальной информации пока не исчерпали свои возможности для получения новых идей. Каналы социальных сетей, использующие режим рейтингования пользователей, являются глобальным источником данных для развития искусственного интеллекта, поскольку позволяют оценить не только поведение пользователей, но и их предпочтения, вкусы, привычки и уровень интеллекта. Обзоры отраслевых данных позволяют получить неструктурированные кросс-канальные данные о состоянии бизнеса, которые ранее не могли быть получены из структурированных источников. Соотношение структурированной и неструктурированной кросс-канальной информации, в конечном счете, формирует негативные и позитивные данные о деятельности компании или конкретном клиенте, что в итоге трансформируется в оценочное суждение об уровне риска и существующих негативных последствиях для бизнеса.

Несмотря на все внимание, которое в последнее время уделяется неструктурированным кросс-канальным данным в ответ на развитие новых технологий оценки и развития искусственного интеллекта, главенствующую роль по-прежнему занимают структурированные данные, обработка и анализ которых менее трудозатратны с точки зрения их использования в практической деятельности.

Ввиду того, что новые финансовые компании, выходя на рынок, демонстрируют все новые способы захвата и анализа неструктурированных данных, Г. Барбье [3] в новом цикле научных исследований был опубликован отчет, в котором подчеркивалась роль структурированных данных в работе финансовых организаций США. В отчете обращается внимание на то, что, несмотря на достигнутые успехи в области сбора, хранения, извлечения и анализа новых форм неструктурированной информации, опрос 300 администраторов баз данных показал высокую заинтересованность большинства кредитных и финансовых компаний в управлении своими структурированными данными, а также планирования продолжения этой работы в обозримом будущем.

Другая часть исследователей [9] занимает противоположную позицию, мнение которых заключается в отсутствии перспектив использования структурированной кросс-канальной информации, поскольку со временем часть ее трансформируется в неструктурированные данные. С такой позицией нельзя не согласиться, учитывая, что развитие новых технологий предопределяет рост неструктурированных данных, объем которых будет продолжать расти, особенно в условиях востребованности новой аналитической информации и инструментов для ее сбора и анализа.

Особенностью кредитно-финансовой сферы является необходимость обеспечения непрерывности бизнес-процессов, поэтому существующая избыточность сети может быть использована для создания большой системы финансовых данных. Некоторые авторы признают [11], что большинство внутренних данных, полученных из внутренних источников, не могут быть использованы в полном объеме. Поэтому применение новых кросс-канальных технологий позволяет получать представление о ранее неиспользованных источниках

неструктурированных данных в электронной почте, записях обслуживания клиентов, данных датчиков и журналах безопасности. В ответ на совершенствование информационных технологий большой интерес представляет поиск новых идей, основанных на анализе больших данных, которые в основном являются внешними по отношению к кредитным и финансовым компаниям, например, социальных сетей, местоположения мобильных телефонов, трафика и погоды.

В международной практике можно найти достаточное количество примеров, когда кредитные и финансовые компании начинают реализовывать собственные конкурентные преимущества на основе использования общего анализа данных. Для многих банков потоки данных в социальных сетях все чаще становятся неотъемлемым компонентом стратегии цифрового маркетинга. На исследовательском этапе эта технология может использоваться для быстрого поиска через огромное количество кросс-канальных потоковых данных и шаблонов существующих трендов, которые относятся к конкретным банковским продуктам или клиентам банков.

Однако для накопления достаточной статистики процесс поиска и накопления необходимо повторять на ежедневной основе, чтобы получить качественный скачок от идентификации шаблона до внедрения нового продукта или услуги для клиента в рамках целенаправленного маркетинга. Кроме того, кодификация взаимосвязи между большой аналитикой кросс-канальных данных и операционными данными требует создания особой модели для ее интеграции.

Большие данные оказывают значительное влияние на многие аспекты управления информацией, включая интеграцию данных.

Традиционно интеграция кросс-канальных информационных ресурсов была сосредоточена на их перемещении через промежуточное программное обеспечение – так называемые «фильтры спецификаций передачи сообщений» для создания нового интерфейса продукта или услуги. Однако эти концепции информационной интеграции в кредитно-финансовой сфере уже устарели, поскольку они адаптированы для управления уже имеющимися информационными данными, оставляя без внимания «подвижную» информацию [12].

Переход современных банков и кредитно-финансовых организаций в новый мир неструктурированных данных и потоковых данных существенно изменил традиционное представление об интеграции ключевых кросс-канальных ресурсов. Существующие передовые технологии сегодня позволяют включить анализ потоков кросс-канальных данных в собственный бизнес-процесс для принятия решения в режиме реального времени. По результатам проведенного анализа можно формализовать общие подходы для

интеграции или включения полученных кросс-канальных данных в существующий банковский бизнес-процесс или бизнес-операции в режиме реального времени.

На основе анализа структурированных и неструктурированных кросс-канальных информационных ресурсов можно получить не только данные о реальной стоимости той или иной компании при выборе контрагентов или заемщиков, но и получить более глубокую информацию о безопасности их бизнеса, что ранее не было возможно ввиду технологических ограничений.

Некоторые специалисты в области информационных технологий сегодня разделяют мнение о том, что развитие новых методов для высокоскоростного перемещения структурированных и неструктурированных кросс-канальных ресурсов является основной задачей для интеграции информации между большими данными и рабочими данными.

В практической деятельности неструктурированные кросс-канальные ресурсы зачастую необходимо быстро перемещать на больших географических расстояниях для совместного использования и совместной работы. Поэтому связывание традиционных информационных источников с большими данными – это многоступенчатый процесс, требующий использования потоковой передачи больших данных и определения соответствующего шаблона. Только после определения объема необходимых данных для управления и анализа необходимо использовать

интеграцию как инструмент для перемещения структурированных и неструктурированных данных.

Основой любой среды информационных данных для компаний и организаций кредитно-финансового сектора выступают новые формы продуктов или услуг, доставляемые до потребителя в режиме реального времени. Именно по этой причине такие данные должны быть масштабируемыми и защищенными от внешнего проникновения. В качестве небольшого отступления лишь отметим, что для цели упорядочивания информации используют четыре наиболее распространенных модели:

- Иерархическая модель, которая ориентирована на описание объектов, находящихся во взаимосвязи. Например, при описании организационной структуры традиционно используется именно иерархическая модель.

- Сетевая модель, которая является более совершенной формой иерархической модели, описывает более сложные виды взаимосвязи между данными или объектами.

- Объектная модель, как новая форма взаимосвязи между отдельными структурами или объектами, иногда именуемыми атрибутами для обработки больших массивов структурированных данных. Например, при ранжировании заемщиков в кредитном портфеле может быть использована объектная модель, связывающая клиентов по фамилии, имени и отчеству, а также номеру кредитной карты. Для этих целей должны быть определены значения указанных атрибутов, что позволяет в дальнейшем выделить объекты, относящиеся к определенному классу.

- Реляционная модель, данные в которой представлены в виде таблиц, состоящих из определенного набора строк. В данном случае строка таблицы является самостоятельным объектом для анализа, а столбцы таблицы выполняют функцию свойства данного объекта. С помощью реляционной модели обеспечивается информационная взаимосвязь между объектом анализа и его функциями. Данная модель является самой распространенной в практике компаний кредитно-финансовой сферы.

Основная причина такой популярности заключается в том, что в основе потребительского предложения банка всегда лежит сочетание возможности использования различных информационных систем, одной из которых является система управления реляционными базами данных (РСУБД), выбор об использовании модели строится в соответствии с уровнем ее производительности, масштаба и стоимости конечного продукта или услуги.

Применение на практике реляционных баз данных сопряжено с необходимостью анализа неструктурированных кросс-канальных информационных источников с помощью программно-технических средств. Использование различных программных инструментов для работы с реляционными базами данных обусловлено типом анализируемых информационных данных или отдельных транзакций. Поэтому различные разработчики баз информационных данных при выборе неструктурированных кросс-канальных информационных ресурсов проводят их тестирование на предмет наличия таких признаков как атомарность, согласованность, изоляция и долговечность.

- Атомарность: при анализе транзакций или информационных данных должен быть получен конкретный ответ на вопрос «все или ничего», только в этом случае можно говорить о наличии признака атомарности. Если какая-либо часть транзакции или базовая кросс-канальная информационная система выходит из строя, вся транзакция завершается с ошибкой.

- Согласованность. В кросс-канальной информационной базе данных всегда должны присутствовать только транзакции с достоверными данными. Если данные повреждены или не соответствуют действительности, транзакция не будет завершена, и данные не будут записаны в информационную базу учетных данных.

- Изоляция. Совершение нескольких одновременных транзакций не должны противоречить друг другу, в противном случае информация будет носить «конфликтный» характер, т.е. противоречить условиям первоначальной записи.

- Долговечность. Данный признак характеризует то, что любая транзакция, записанная в базу, должна быть сохранена и оставаться на постоянной основе.

Некоторые специалисты отмечают, что высокие требования, предъявляемые к неструктурированным кросс-канальным информационным ресурсам, предопределены необходимостью их последующей виртуализации, т.е. отделения информации от базовой физической среды, ее последующей обработке и доставке в рамках единой кросс-канальной информационной системы.

В кредитно-финансовой сфере одной из основных причин внедрения технологии виртуализации информационных данных является высокая скорость обработки информации и внедрение большого количества операционных систем. Именно поэтому важнейшим условием для полноценной обработки структурированных и неструктурированных кросс-канальных информационных ресурсов является внедрение технологии гипервизора.

В научной и деловой литературе, посвященной внедрению технологий гипервизора, все чаще звучат мнения о необходимости его использования в практике кредитно-финансовой деятельности компаний и организаций, осуществляющих деятельность на открытом финансовом рынке в целях повышения производительности и эффективности обработки разнообразных структурированных и неструктурированных кросс-канальных информационных ресурсов.

С научной точки зрения, гипервизор представляет собой новую технологию, обеспечивающую совместное использование кросс-канальных информационных ресурсов на упорядоченной основе. Кроме того, практическое применение гипервизора обеспечивает динамическое распределение информации в глобальной операционной системе.

Например, кросс-канальные информационные ресурсы могут быть разделены пополам между материнской и сторонней операционной системой, что затрудняет обработку информации, делая ее сложной и неэффективной. С помощью гипервизора решение этой задачи значительно упрощается путем компоновки и распределения между соответствующими информационными массивами данных.

В отличие от структурированных кросс-канальных информационных ресурсов, неструктурированные ресурсы состоят из данных, собранных из разных источников внутри кредитно-финансовой организации, поэтому необходимо также учитывать затраты, связанные с очисткой и обработкой этих данных, поскольку:

- Традиционные потоки кросс-канальных данных, полученных из транзакций, приложений и пр. информационных источников могут создавать множество несопоставимых между собой значений.

- Существуют также десятки новых кросс-канальных информационных ресурсов, каждый из которых нуждается в определенной структуризации, прежде чем его можно будет использовать для практической работы.

- Неструктурированные кросс-канальные информационные ресурсы должны быть обработаны и соответствовать формату структурированных ресурсов, что требует использования различных методов и технологий.

Как указывают некоторые авторы [5], исторически сложилось так, что содержимое хранилищ информационных данных в крупных кредитно-финансовых организациях скомпоновано таким образом, чтобы предоставлять возможность в режиме реального времени доставлять любую информацию для решения таких функциональных проблем как поддержка клиентов, продажа и услуг клиентам, обслуживание в ближайшем реальном времени. Таким образом, хорошо структурированное хранилище данных помогает оперативно перемещать структурированные и неструктурированные кросс-канальные информационные ресурсы из бэк-офиса напрямую к клиенту.

Алгоритмы работы с неструктурированными кросс-канальными информационными ресурсами имеют много особенностей. В частности, они должны включать в себя ряд инструментов, обеспечивающих их формализацию и консолидацию в режиме реального времени из различных источников, что предопределяет наличие соответствующей эталонной архитектуры.

В научной литературе встречаются различные мнения относительно использования эталонной архитектуры для работы с неструктурированными кросс-канальными

информационными ресурсами. Однако в большинстве источников [13] выделяют как минимум три фундаментальных инструмента (отчеты и информационные панели, режим визуализации данных, расширенная аналитическая платформа).

- Отчеты и информационные панели. Практическое использование этого инструмента обеспечивают «удобное» представление информации из различных информационных источников. Несмотря на то, что в традиционном банкинге сегодня активно развиваются дистанционные технологии для оперативного доступа клиентов к различным информационным ресурсам, развитие информационных отчетов по-прежнему актуально и для крупных и для малых кредитно-финансовых организаций.

- Режим визуализации данных. Использование данного инструмента определяет дальнейшие шаги в развитии финансовой отчетности. Чем более интерактивным и динамичным будет режим визуализации данных, тем эффективнее будет построена работа по взаимодействию с клиентами. Сегодня пользователи любой бизнес информации могут наблюдать за динамичным изменением данных в режиме реального времени, используя различные методы визуализации, включая инфографику, диаграммы, видео контент и пр., что определяет режим их деловой активности.

- Расширенная аналитическая платформа. Так называемая «умная аналитика» является новым инструментом, определяющим ключевые тенденции или события в кредитно-финансовом секторе. Ее использование позволяет преобразовывать традиционную информацию в уникальные аналитические данные, востребованные клиентами в деловой практике. Например, в крупнейших кредитных организациях сегодня повсеместно внедряются продукты «умной аналитики», включающие возможность построения прогнозных инвестиционных и бизнес моделей для компаний активно реализующих крупные инфраструктурные проекты. Уникальным примером «умной аналитики» также является возможность создания индивидуальных бизнес планов для физических лиц, размещающих свои средства в депозиты и сберегательные сертификаты.

В мировой практике информатизации кредитно-финансовой сферы все большее распространение получает создание общественных и частных облачных ресурсов.

В качестве небольшого отступления лишь отметим, что в научной и деловой литературе термин «облачные вычисления» появился относительно недавно. Развитие данной инфраструктуры было предопределено необходимостью оперативного обеспечения удобного сетевого доступа для пользователей к определенному пулу конфигурируемых и вычислительных ресурсов (например, мобильного приложения, сервисам связи и пр.), которые могут быть оперативно предоставлены пользователю с минимальными эксплуатационными затратами и обращениями к провайдеру этих продуктов и услуг.

В практической деятельности кредитно-финансовые организации традиционно используют несколько типов облачных вычислений общественные, публичные или частные.

Создание общественного облака как нового вида информационной инфраструктуры для организаций кредитно-финансового сектора обусловлено в первую очередь необходимостью обработки большого количества неструктурированных данных при условии обеспечения безопасности и конфиденциальности их дальнейшего применения. Согласно установленным международным принципам, общественное облако может находиться в совместной собственности одной или нескольких организаций и эксплуатироваться на паритетной основе, что позволяет снижать затраты на его содержание.

Публичное облако представляет собой набор аппаратных средств, сетей, хранилищ, сервисов, приложений и интерфейсов, которыми владеет и управляет третье лицо для использования другими компаниями и отдельными лицами. Эти коммерческие поставщики создают масштабируемый центр обработки данных, который скрывает детали базовой инфраструктуры от потребителя. Примерами создания публичного облака являются специализированные банковские информационные сети, почтовые ресурсы и приложения для клиентов. Поставщик облачных вычислений может оптимизировать информационную среду для организации поддержки большого числа клиентов кредитно-финансовой организации в любой точке мира.

Частное облако – представляет собой определенный набор аппаратных средств, сетей, хранилищ, сервисов, приложений и интерфейсов, которыми владеет и управляет кредитно-финансовая организация для использования ее сотрудниками, партнерами и клиентами. Особенностью частного облака является то, что оно может быть создано и управляться третьей стороной только в целях одной кредитно-финансовой организации. Являясь частью высокоавтоматизированной инфраструктуры, частное облако позволяет заменить ручные процессы управления информационными услугами для поддержки клиентов; использовать большие массивы неструктурированных кросс-канальных информационных ресурсов и управлять ими в условиях обработки огромных объемов данных.

Необходимо принимать во внимание, что традиционные инструменты для интеграции информации получают новый импульс развития ввиду необходимости обработки разнообразных неструктурированных данных и растущего объема и скорости их передачи конечному пользователю, в то время как традиционные формы интеграции информации нуждаются в постоянном совершенствовании используемых технологий, что предопределяет снижение объема их использования и переход на более эффективные и менее затратные платформы.

Таким образом, развитие новых технологий обработки, хранения и передачи информации в кредитно-финансовой сфере предопределено не только существенными различиями между структурированными и неструктурированными информационными ресурсами, но и эффективностью их использования в условиях активного применения дистанционных банковских продуктов и услуг. Всеобщая глобализация кредитно-финансовой сферы и переход на новый уровень технологий диктует качественно иную стратегию развития информационных ресурсов при соблюдении следующих принципов:

- Понимания способов анализа информационных данных во взаимосвязи сопоставимых и несопоставимых параметров, используемых аналитических инструментов и качества предоставления информации клиенту.

- Создания новой инфраструктуры для обработки информации. Обработка и анализ неструктурированных кросс-канальных информационных ресурсов имеют высокие операционные затраты, что определяет поиск новых направлений для ее интеграции и последующей доставке пользователю.

- Оптимизация существующих способов интеграции информационных данных в условиях растущего объема программно-технических средств, технологий и операционных систем для обеспечения конкурентного преимущества кредитно-финансовой организации.

С нашей точки зрения, реализация данных принципов будет способствовать переходу кредитно-финансовой сферы на новый уровень развития в рамках создания новых структур управления и обработки информации, а также решения стратегических задач по обеспечению безопасности хранения и передачи конфиденциальных данных по операциям клиентов и контрагентов.

**Исследование выполнено при финансовой поддержке РФФИ  
в рамках научного проекта № 20-010-00101 А.**

## ЛІТЕРАТУРА

1. Amir Gandomi and Murtaza Haider, «Beyond the hype: Big data concepts, methods, and analytics,» *International Journal of Information Management*, vol. 35, Issue 2, pp 137-144, April 2015.
2. Blechar M., Adrian M., Friedman T., Schulte W. R. and Laney D., “Predicts 2012: Information Infrastructure and Big Data,” Gartner Inc., November 29, 2011;
3. Barbier G., and Liu H., «Data mining in social media», C. C. Aggarwal (Ed.), *Social network data analytics*, Springer, pp. 327–352, 2011.
4. Charu C. Aggarwal, «An Introduction to Social Network Data Analysis,» *Social Network Data Analytics*, Springer, 2011.
5. Dener, C., Watkins, J., Dorotinsky, W. L. (2011), «Financial Management Information Systems: 25 Years of World Bank Experience on What Works and What Doesn't», World Bank Publications, Washington.
6. Gantz J. and Reinsel D., «Extracting value from chaos,»' in *Proc. IDC iView*, pp. 1-12, 2011.
7. Heidemann J., Klier M., and Probst F., «Online social networks: A survey of a global phenomenon», *Computer Networks*, vol. 56(18), pp. 3866–3878, 2012.
8. Manyika J., Chui M., Brown B., Bughin J., Dobbs R., Roxburgh C., et al., «Big data: The next frontier for innovation, competition, and productivity,» McKinsey Global Institute Article, May 2011.
9. Parthasarathy S., Ruan Y., and Satuluri V., «Community discovery in social networks: Applications, methods and emerging trends,» C. C. Aggarwal (Ed.), *Social network data analytics*, Springer, pp. 79–113, 2011;
10. Randy Rieland, «Big Data or Too Much Information», *Smithsonian Magazine*, May 7, 2012.
11. Reena Duggal, Balvinder Shukla and Sunil Kumar Khatri, «Big Data Analytics in Indian Healthcare System – Opportunities and Challenges» research paper accepted at National Conference on Computing, Communication and Information Processing (NCCCIP-2015) – May 2015: 92-104.
12. Rupanagunta K., Zakkam D., and Rao H., «How to Mine Unstructured Data», *Article in Information Management*, June 29, 2012;
13. Van der Heijden, H., Van der Heijden, J. G. M. (2009), «Designing Management Information Systems», Oxford University Press, Oxford.;
14. Wijnhoven, F. (2009), «Information Management: An Informing Approach», Routledge, New York.