



**Экономический
и Социальный Совет**

Distr.: General
2 April 2015
Russian
Original: English

Европейская экономическая комиссия

Конференция европейских статистиков

Шестдесят третья пленарная сессия

Женева, 15–17 июня 2015 года

Пункт 4 предварительной повестки дня

**Модернизация статистического производства и услуг
и управление, ориентированное на эффективность**

Растущее значение "оптовой поставки" данных

Записка Австралийского бюро статистики

Резюме

В настоящем документе описана существующая в Австралийском бюро статистики стратегия распространения данных, в которой признается, что присутствие в сети более не определяется основным веб-сайтом национального статистического управления, которое должно отныне заявлять о себе и гибко действовать через различные каналы. В документе рассматривается значение понятия "оптовой поставки" данных, а также способы получения положительных результатов Австралийским бюро статистики и многими другими организациями за счет использования возможности взаимодействия с хранилищами данных официальной статистики для того, чтобы закачивать эти данные в их приложения.

Документ представлен для обсуждения на втором заседании семинара Конференции европейских статистиков "Модернизация статистического производства и услуг и управление, ориентированное на эффективность".

GE.15-07038 (R) 190515 200515



* 1 5 0 7 0 3 8 *

Просьба отправить на вторичную переработку



I. Справочная информация

1. Одной из законодательно закрепленных функций Австралийского бюро статистики (АБС) является максимально широкое использование данных для обоснованного принятия решений. АБС действует внутри информационной среды, состоящей из взаимосвязанных элементов, и играет ведущую роль в обеспечении Австралии статистической информацией на основе эффективных партнерских связей с ключевыми заинтересованными сторонами. В результате изменений и инноваций в информационной среде растет спрос на быстрый, гибкий и безопасный доступ к данным со стороны государственных органов, исследовательских организаций, предприятий и общества в целом.

2. В настоящем документе изложена стратегия повышения эффективности доступа к статистической информации путем "оптовой поставки" данных. В нем подробно рассматриваются положительные аспекты использования АБС, другими организациями и частными лицами возможности подключения к хранилищам данных АБС для зачивания из них нашего контента в их приложения через веб-сервисы или интерфейсы прикладного программирования (ИПП).

II. Наши пользователи

3. Традиционный подход АБС к своим пользователям основан на сегментации рынка, начиная с учащихся начальной школы и заканчивая специалистами по анализу политики, созданию экономических моделей и исследователями в области здравоохранения. С недавнего времени АБС переключило свое внимание с организации, к которой принадлежат пользователи, или роли, которую они играют, на их поведение. Это позволило нам упростить наш подход к предоставлению данных.

4. Пользователей можно разделить на четыре основные поведенческие группы, что позволяет нам формировать различные подходы для удовлетворения их потребностей. Что говорят эти пользователи:

- мне нужна цифра, которую я могу быстро использовать, передавая эту информацию другим;
- мне нужна интересная история, из которой я узнаю что-нибудь новое; мне хотелось бы услышать историю, увидеть интерактивный инструмент или игру, например такую, как "В центре внимания – перепись" (spotlight.abs.gov.au);
- мне нужна краткая информация, так как я хотел бы узнать больше о том месте, где я живу, или о важных проблемах здравоохранения сегодняшнего дня;
- мне нужны подробные данные, которые я смогу анализировать, повторно использовать и сочетать с другими данными, что поможет мне или моим собственным пользователям в принятии решений.

III. Цифровая экономика

5. В течение последних десятилетий организации в своей работе опирались на статические веб-сайты, браузеры для настольных компьютеров и монолит-

ные системы управления контентом. В нашей традиционной модели используется веб-сайт, на котором в заранее определенные и прогнозируемые моменты регулярно появляются статические публикации и данные. Пользователи имеют возможность получить подробные данные в формате Excel, однако сделать это они могут, только загрузив их вручную с веб-сайта.

6. В результате стремительного технологического прогресса эта модель утратила свою эффективность. Сложный характер быстрого развития технологий, возможностей обработки информации, цифровых средств информации и цифрового сообщества приводит к появлению ряда вызовов. Нам известно следующее:

- объем данных растет экспоненциально, но при этом стоимость их хранения или передачи сильно уменьшается. В результате этого появляются затруднения в управлении данными и их распространении;
- основная масса добавленной стоимости теперь создается не провайдерами, а агрегаторами или посредниками в области данных, такими как новые медийные платформы. Задача состоит в том, как привлечь внимание или обеспечить атрибуцию (а тем самым и качество) в ситуации, когда создание контента и каналы распространения не ограничены традиционными средствами;
- Способность взаимодействовать с новыми медийными средствами выходит за рамки традиционных жестких моделей. Пользователи могут получить мгновенный доступ к контенту, обмениваться им, комментировать его и делать запросы в любое время суток.
- скорость в цифровой экономике определяет все. Миллисекунды – это деньги. Все хотят, чтобы информация была под рукой, а получать ее можно было быстро и эффективно;
- важным игроком является сообщество разработчиков приложений. Благодаря новым веб-технологиям новые приложения можно разрабатывать быстро и с небольшими затратами для удовлетворения широких или конкретных потребностей;
- информация может быстро и автоматически толковаться с помощью машин, т.е. машины способны выполнять большую часть утомительной работы в сети по поиску, компиляции информации и дальнейшим действиям на ее основе;
- в Интернете все более широко используются системы искусственного интеллекта, что позволяет автоматизировать процесс принятия решений.

7. Наши основные пользователи являются действующими лицами цифровой революции и испытывают на себе ее влияние. Государственным органам, деловым и научно-исследовательским кругам необходим более быстрый и улучшенный доступ к данным. Им нужен доступ к подробным, полностью описанным макроданным и обезличенным микроданным. Для решения сложных проблем им также нужна возможность интегрировать данные по традиционным и новым тематическим областям статистического учета, почерпнутые из различных источников.

8. Они не хотят иметь дело с традиционной моделью. Загрузка файлов данных с веб-сайта вручную – неэффективный, медленный и непродуктивный процесс, который препятствует инновациям. Результат нашей работы – это зачастую важные исходные данные для их бизнес-процессов.

9. Некоторые пользователи ждут от нас предоставления данных на основе согласованных определений структуры данных в таких форматах, как Стандарт обмена статистическими данными и метаданными (SDMX). Это эффективно для обеих сторон и позволяет им моментально извлечь нашу информацию или получить к ней доступ. В конечном счете согласование стандартов в области метаданных и данных в статистической отрасли упростит и упорядочит пользователям доступ к множеству источников данных, вне зависимости от источника.

IV. Трансформация результатов работы и коммуникации

10. При разработке нового подхода к распространению данных АБС ориентируется на следующие принципы и ставит перед собой следующие цели:

- обеспечить более высокую степень согласованности благодаря хранению наиболее детализированных и наиболее полно описанных многомерных данных в хранилище итоговых данных, которое будет являться "единственным источником истины";
- поощрять и использовать стандартные подходы к метаданным не только внутри самой организации, но и во всей статистической и информационной отрасли;
- сотрудничать с организациями, занимающимися аналогичной работой, для обмена информацией о подходах и инструментах;
- компоновать множество вариантов потенциальных итоговых данных на базе "единственного источника истины";
- динамически компилировать контент, чтобы удовлетворить потребности пользователей, имеющих различные уровни статистической грамотности; хотя некоторый объем статического контента будет сохранен, все чаще контент будет формироваться динамически на основе детализированных базовых данных и метаданных;
- представлять данные и метаданные на основе стандартной библиотеки инструментов с помощью географических карт, графиков и инструментов визуализации данных;
- предоставлять свободный, своевременный и не сопряженный с большими расходами доступ к информации государственного сектора;
- обеспечивать доступ к данным, в том числе к их большим массивам, используя сервисы типа "машина-машина";
- сосредоточить внимание на потребностях пользователей, своевременно предоставляя им качественные и актуальные статистические данные таким образом, чтобы пользователи могли легко находить и понимать их;
- обеспечивать предоставление необходимых данных независимо от того, через какой канал или устройство их получают пользователи;
- содействовать повторному использованию данных АБС внешними пользователями и принадлежащими третьей стороне порталами, обеспечивая при этом узнаваемость товарного знака АБС и возможность определения АБС как ее источника;

- обеспечивать для контента, предназначенного для укрепления статистического потенциала, заметное место наравне с данными, но не препятствующее их восприятию.

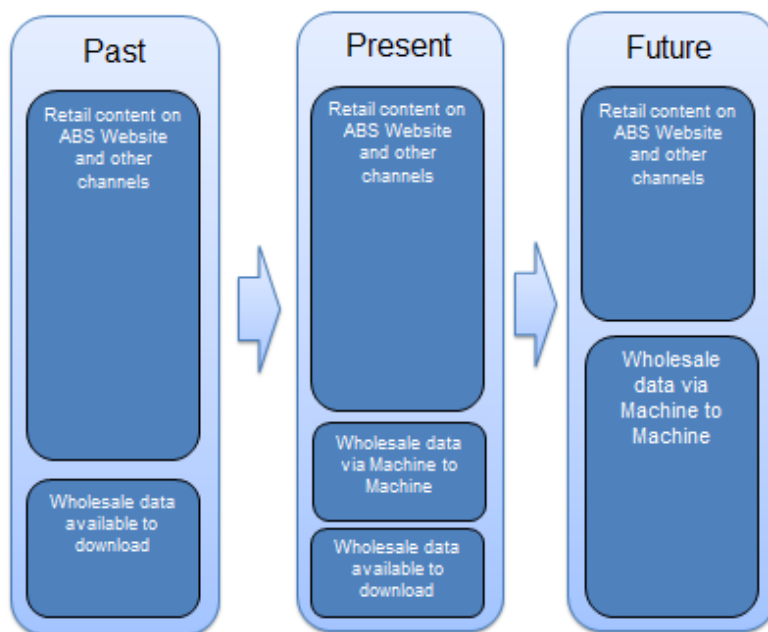
11. Один из важных элементов реагирования на цифровую революцию – "оптовая поставка" наших данных. Под "оптовой поставкой" мы подразумеваем предоставление доступа к большим объемам данных/контента с помощью сервисов "машина–машина", с тем чтобы другие пользователи могли далее их использовать в процессе формирования более ценных продуктов или услуг для себя или для своих пользователей.

12. Это не означает, что АБС больше не будет само предоставлять данные, контент или приложения. Мы несем серьезные обязательства перед обществом и всегда будем предоставлять австралийской общественности полные данные по широкому кругу вопросов и краткий контент. АБС способно удовлетворить существующий со стороны пользователей спрос на надежные, комплексные статистические решения.

13. Вместе с тем это все же означает, что мы будем все более активно задействовать в собственной работе и предоставлять другим решения типа "машина–машина" для интеграции, агрегирования и вторичного использования наших более детализированных данных в форматах и в рамках инновационных решений, которые само АБС не предоставляет. Такой подход позволяет другим использовать наши данные, ориентируясь на свои конкретные цели, в том числе использовать их в своих собственных системах и рабочих процессах, включать их в широко распространенные коммерческие аналитические инструменты и использовать их для оказания услуг с добавленной стоимостью.

Рис. 1

Трансформация результатов работы АБС



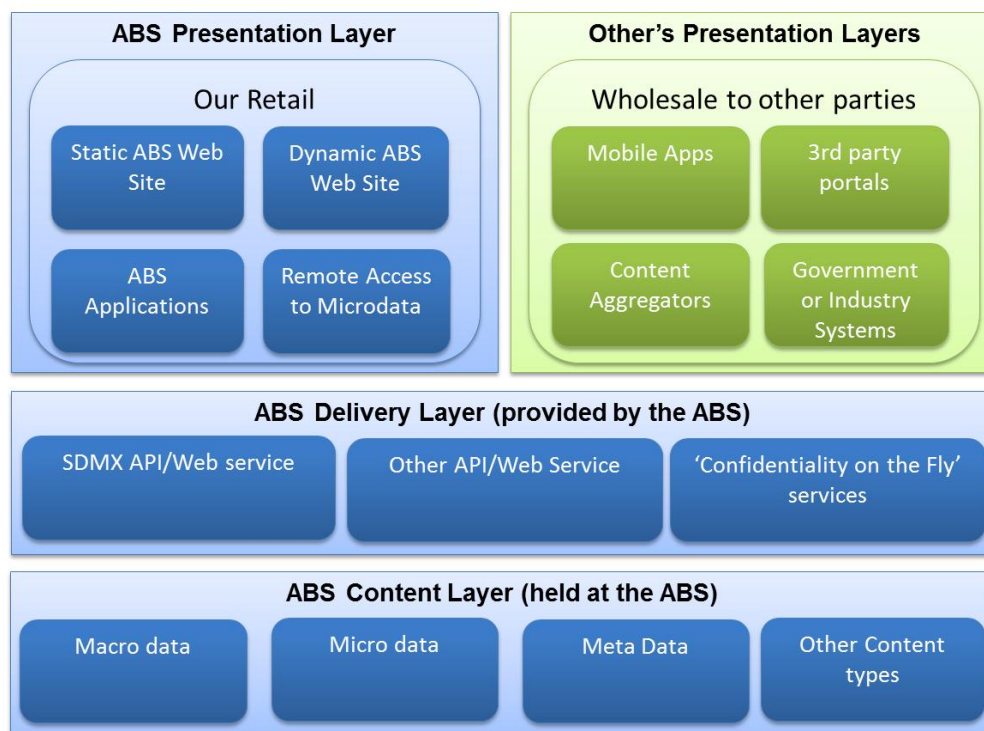
1

A. Гибкая архитектура для динамичной публикации

14. В силу такого подхода необходима архитектура (в том числе бизнес-архитектура, системная и информационная архитектура), которая предусматривала бы отдельные слои для контента, его поставки и представления.

Рис. 2

Трехуровневая модель для поддержки "оптовой поставки" данных



15. АБС должно иметь возможность строить структуру, каталоги, создавать описания, хранить, защищать и предоставлять данные систематическим образом вне зависимости от того, как они впоследствии будут использоваться или потребляться нашими пользователями. Вместо создания традиционных публикаций, где сводные данные, детализированные данные, анализ и примеры, касающиеся данных, могут существовать в рамках одной монолитной и статической публикации, мы сосредоточим свои усилия на создании базовых элементов контента.

16. Слой поставки служит связующим звеном между контентом и его представлением. Важно отметить два аспекта, касающихся слоя поставки. Первый относится к инструментам, которые позволяют пользователям и системам получать доступ к контенту различной степени детализации, а второй – к сервисам, известным как ИПП или веб-сервисы, благодаря которым данные самой высокой степени детализации могут быть преобразованы и обезличены, прежде чем они станут доступны нашим пользователям ("динамическое обеспечение конфиденциальности"). Такие сервисы, при условии использования соответствующих метаданных, позволяют делить контент самым разным образом и объединять его с контентом из каналов, принадлежащих третьим сторонам.

17. Презентационный слой дает возможность эффективно удовлетворять различные потребности пользователей благодаря тому, что мы можем представлять контент разными способами, в разных форматах и на разных устройствах. Презентационный слой охватывает не только розничное предоставление статистических данных и контента по каналам АБС, но и представление данных и контента, которые АБС оптом поставило другим сторонам. В него будет включаться все больше общих и специализированных приложений, которые предоставляем мы сами или другие.

V. "Оптовая поставка" данных для собственного использования

18. Примером "оптовой поставки" данных для собственного использования может служить приложение АБС "Данные по регионам" (DbR) (www.abs.gov.au/databyregion).

19. Пользователи хотят получать информацию о тех местах, где они живут. Они хотят сравнивать свой регион с другими регионами и понимать его положение в более широком национальном контексте. В течение последнего десятилетия одним из механизмов предоставления такой информации являлся инструмент АБС "Обзор национальных регионов" (ОНР). Благодаря ОНР данные по регионам из многочисленных наборов данных представлены в виде единого последовательного динамического ряда. В прошлом, для того чтобы получить ОНР, нам приходилось вручную верстать и выверять тысячи отдельных веб-страниц, посвященных регионам. Это избавляло пользователей от необходимости искать и изучать каждый набор данных самостоятельно, однако сам процесс занимал месяцы, был ресурсоемким и неэффективным. После публикации данные ОНР обновлялись только в момент следующей ежегодной публикации, несмотря на постоянное обновление данных о регионах в течение года.

20. В июле 2014 года АБС разработало приложение "Данные по регионам". Чтобы предоставить нашим пользователям новую, инновационную услугу, в этом приложении мы задействовали все три слоя:

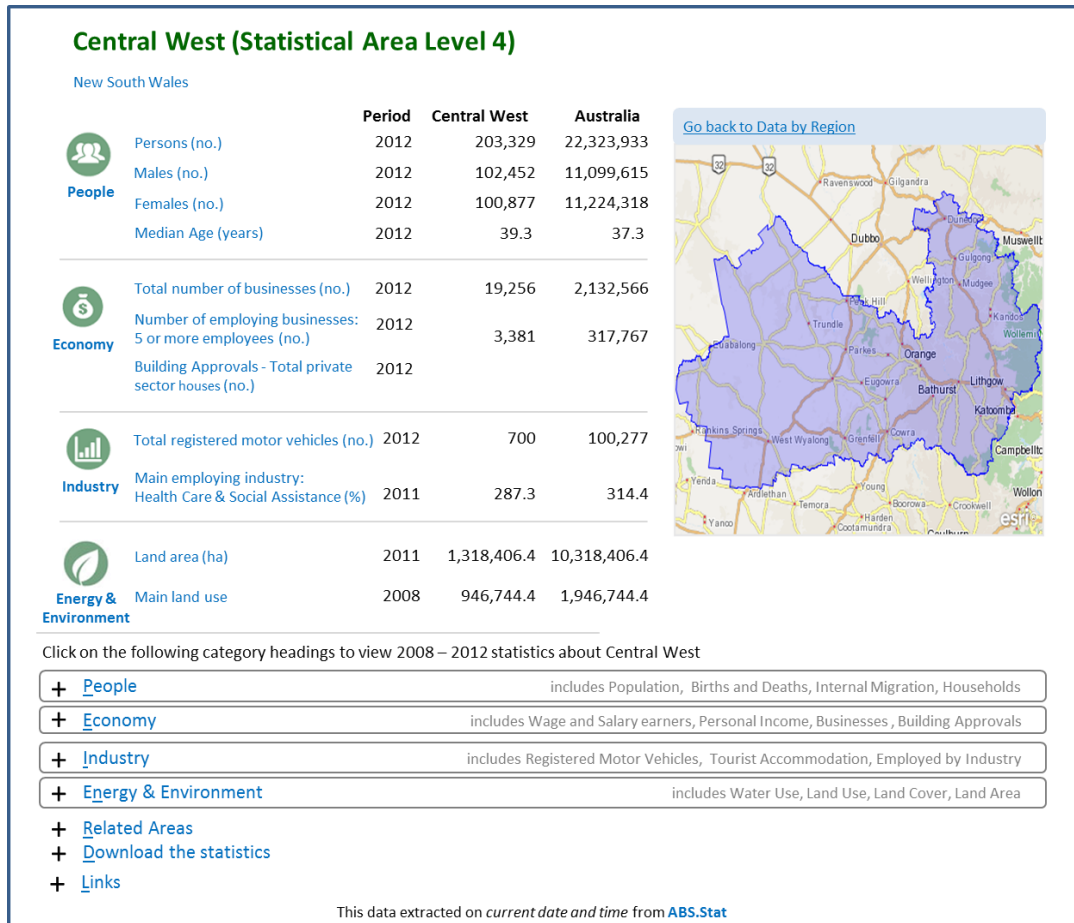
- слой контента формируется через наборы данных из базы ABS.Stat (в ее основе лежит статистическая база данных ОЭСР OECD.Stat, которая используется в ряде организаций официальной статистики);
- слой поставки формируется через веб-сервисы ABS.Stat и веб-сервис, разработанный АБС; и
- презентационный слой, формируется приложением "Данные по регионам".

21. Приложение "Данные по регионам" является единой точкой доступа к статистическим данным АБС в разбивке по регионам. В нем предусмотрен интуитивно понятный режим навигации при помощи картографического инструмента, а новые статистические данные доступны по мере их появления. Самое важное революционное изменение в рамках проекта "Данные по регионам" не лежит на поверхности; оно касается способа вызова информации/данных из единого централизованного хранилища данных. Когда пользователь выбирает для просмотра тот или иной регион, данные динамически извлекаются из хранилища. У нас больше нет необходимости создавать отдельные наборы данных для каждого региона вручную. Благодаря этому подходу наши собственные данные оптово используются для наших собственных розничных приложений.

Это поможет нам выстроить устойчивый и гибкий подход к распространению нашей информации.

Рис. 3

Данные по Центрально-Западному региону Нового Южного Уэльса (НЮУ)
(из приложения "Данные по регионам")



VI. Другие пользователи "оптовой поставки" наших данных

22. Новые веб-сервисы АБС позволяют также другим сторонам в реальном времени получать доступ к данным, безопасно хранящимся в АБС, интегрировать эти данные с другими наборами данных и представлять результат в своих собственных системах. В качестве примера можно привести Австралийскую сеть исследовательской инфраструктуры по вопросам городов (AURIN), которая действует во всем университетском секторе Австралии и за его пределами (см. <http://aurin.org.au/>).

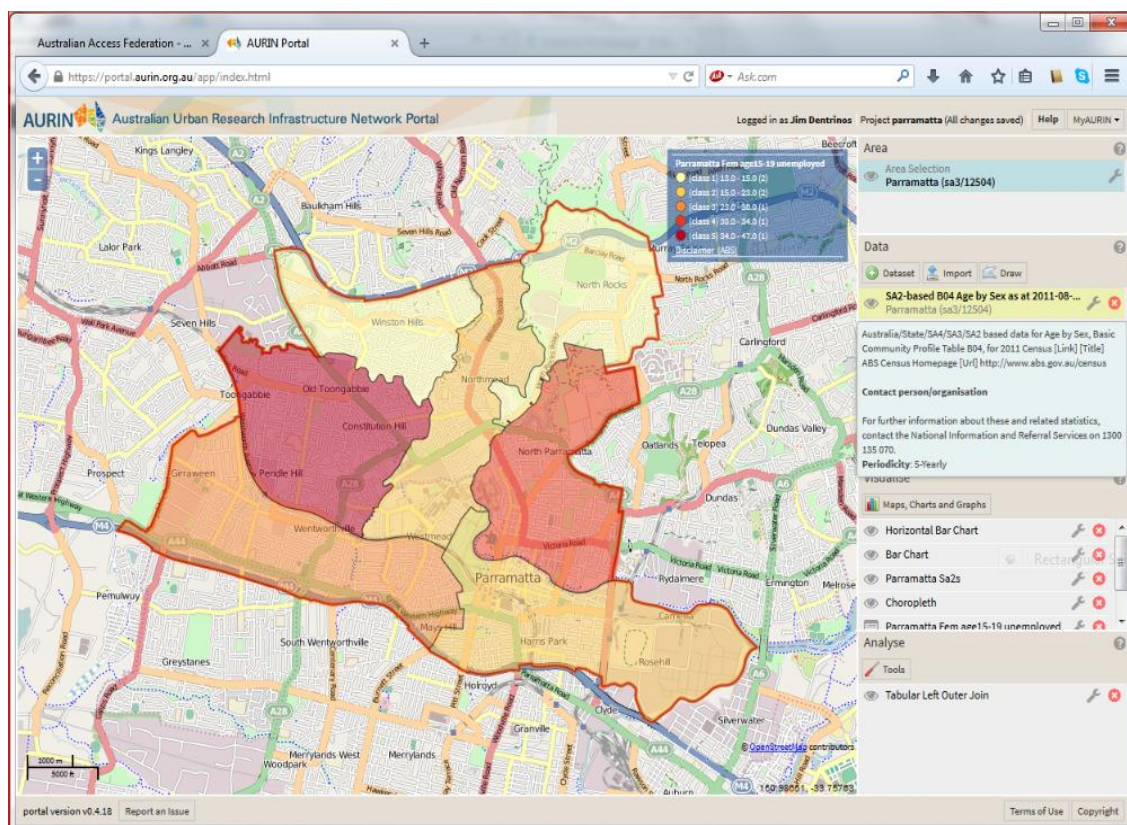
23. Задача AURIN – содействовать исследованиям городских поселений и принятию соответствующих решений путем развития электронной инфраструктуры, которая является двигателем и предвестником инноваций, новых методов и практики. В середине 2013 года AURIN и АБС приступили к осуществлению совместного проекта по предоставлению AURIN доступа к наборам данных

АБС через работающие онлайн и в режиме реального времени сервисы "машина-машина".

24. Этот проект позволил успешно продемонстрировать, как статистические данные передаются по всей территории Австралии в режиме "машина-машина" на примере доступа к данным переписи 2011 года в модели концентратора федерированных данных. Участники AURIN задействовали имеющиеся у них знания и инструменты, чтобы представить эти данные через сложный пользовательский интерфейс, который АБС никогда не смогло бы разработать. По сути, АБС осуществило "оптовую поставку" данных AURIN, которая создала новые продукт и услугу с добавленной стоимостью (портал AURIN).

Рис. 4

Снимок экрана портала AURIN, где используются данные АБС



25. "Оптовая поставка" данных АБС дала ряд преимуществ как AURIN, так и самому АБС, в том числе:

- представляемые AURIN данные – это всегда новейшие данные АБС;
- возможность повторного использования услуг: портал AURIN касается вопросов городской инфраструктуры, но наши технологии веб-сервисов вполне могли бы применяться и для создания портала по вопросам экологии или экономики;
- повышение производительности пользователей AURIN;
- АБС не пришлось разрабатывать презентационный слой;

- множество различных разработчиков предоставили аналитические методы (как существующие, так и новые) для использования при работе с порталом;
- выросла потребность в хорошо описанных данных и метаданных;
- протестированы некоторые аспекты нашей будущей стратегии распространения данных;
- накоплен дополнительный опыт работы с растущим сектором информационных посредников.

VII. "Оптовая поставка" микроданных

26. Последняя область, где мы добились успеха, – это динамическое формирование агрегированных данных из обезличенных единичных записей. Мы вложили средства в разработку отдельного пакета продуктов, которые позволяют внешним пользователям – исследователям или специалистам по вопросам государственной политики – получать доступ к данным по подгруппам населения без какого бы то ни было ущерба для неприкосновенности личной жизни и конфиденциальности. Эти продукты позволяют пользователям табулировать и анализировать микроданные, а также формировать динамически обезличенные итоговые данные.

27. Примером такого веб-продукта является приложение TableBuilder (см. TableBuilder для обследований). Слой поставки этого приложения работает, в том числе, по методу создания помех для автоматической защиты таблиц дискретных данных и ключевой сводной статистики (такой как установленные классы величины, итоги, средние, медианы и квинтили). Данный метод был разработан в целях снижения риска раскрытия при обработке запросов на схожие таблицы, повторных запросов на идентичные таблицы и повторных запросов на одни и те же ячейки таблицы в различных таблицах. В приложение TableBuilder включены и другие алгоритмы обезличивания данных и процедуры защиты конфиденциальности, которые работают в динамическом режиме и снижают до минимума утрату полезности, одновременно уменьшая вероятность идентификации респондентов по полученным данным.

28. В настоящее время АБС предоставляет презентационный слой для сервисов по обезличиванию данных, а аналитические методы ограничиваются теми, которые мы имеем возможность предоставить. В будущем, однако, мы планируем предоставлять возможность моментального обезличивания данных в качестве отдельной услуги, с тем чтобы другие лица могли формировать обезличенные данные для использования в своих системах. Вместе с подходами по предоставлению несистемных решений доступа к микроданным, такие услуги по моментальному обезличиванию откроют новые возможности для наших пользователей, в особенности исследователей, в части взаимодействия с нашими более детализированными наборами данных и повторного использования.

VIII. В перспективе

29. АБС осознает, что пользователь в сети более не является всего лишь "читателем". Сегодня веб-пользователь – это одно из главных действующих лиц сети. Пользователи все чаще играют и хотят играть активную роль в распространении, обмене, обсуждении, популяризации и повторном использовании

информации. "Оптовая поставка" данных открывает широкий спектр возможностей для интеграции в цифровую экономику. Имея доступ к оптовым объемам данных и метаданных, пользователи смогут извлекать контент АБС и загружать его в предпочтительные для них потоки контента разными способами: через новые медиа, свои собственные системы или инновационные продукты, о которых сегодня пока еще ничего неизвестно. Хотя пользователи по-прежнему будут иметь потребность в контенте АБС, по нашим прогнозам, спрос на "оптовую поставку" данных будет расти.

30. Мы являемся экспертами в области сбора, обработки и анализа статистических данных. Мы не всегда обладаем экспертными знаниями, необходимыми для представления таких данных, и не имеем ресурсов, которые позволили бы нам предоставлять данные во всех форматах или в соответствии с креативными стилями, которых ожидают пользователи. Для этого нам нужны другие участники.

31. Существует ряд вызовов:

- с активизацией повторного использования, сбора в одном месте данных из разных источников и распространения данных появляется необходимость в разработке сложных методов, которые позволят обеспечить атрибуцию данных, точно определять их происхождение и защитить наш товарный знак;
- чтобы добиться стабильного обеспечения надежности и эффективности, наши процессы и инфраструктуру необходимо масштабировать;
- новые подходы являются источником не только новых возможностей, но и вызовов, и мы должны обеспечить их увязку с нашими стратегическими целями и операционной моделью, а также жизнеспособность нашей организации.

32. Вне зависимости от того, делает ли АБС большие или малые шаги, оно твердо уверено, что эту трансформацию провести необходимо, и признает, что многие другие национальные статистические управления и международные учреждения вносят в свою работу аналогичные изменения. "Оптовая поставка" данных является важной частью модернизации и максимального расширения использования данных АБС.

Справочная литература

Branson, M 2014, Measuring Interest Group Expectations and Trust, paper presented to 2014 IAOS Conference, Da Nang, Vietnam, 8-10 October 2014.

Borowik, J., Branson, M. and Watson, D. 2013, Interactive Content and Dynamic Publishing – A vital part of an NSO'S output and communications strategy, paper presented at IASE, Macau 2013.

Montgomery, T 2014, Dynamic Publishing – how far can we go? , paper presented at 2014 IMAODBC Conference, Evora, Portugal, September 2014