

Положительная сторона «БОЛЬШИХ ДАННЫХ»

Кристофер У. Сурдак и Сара Агарвал

ФЕРМЕР пользуется простым мобильным телефоном, чтобы узнать цены на выращенные им культуры, а затем находит покупателя, готового заплатить самую высокую цену: это наиболее распространенный пример того, как технология может стимулировать социально-экономическое развитие.

Но современные технологии позволяют нам перевести этот пример на ступеньку выше. С помощью смартфона и простых приложений покупатели могут отслеживать число фермеров, с которыми они встречаются за день, а фермеры могут следить за своими годовыми доходами. Располагая собранной за год информацией, приложение может подсказать фермерам оптимальные сроки и место для сбыта их продукции. Приложения увязывают рыночный спрос с метеорологическими явлениями, чтобы информировать фермеров об оптимальных сроках посева и уборки урожая. Такие аналитические инструменты также позволяют фермерам узнавать об осадках и почвенных условиях и их влиянии на посеянные ими культуры. Они также могут оповещать фермеров о распространении заболеваний сельскохозяйственных культур и методах их профилактики — щелчком кнопки. Благодаря использованию «больших данных», мелкие фермеры могут принимать более обоснованные решения, что ведет к дальнейшему повышению эффективности и рентабельности.

«Большие данные» часто определяются как сбор, анализ и использование больших объемов данных. Но все это уже делается с помощью существующих инструментов; «большие данные» отличаются от них способностью выявлять незамеченные ранее отношения, пользуясь многочисленными и порой не связанными друг с другом источниками. Мы можем теперь задавать новые вопросы и получать новые ответы. И можем это делать, пользуясь общей инфраструктурой (такой как «облачные» вычисления), а не установленным на месте оборудованием, таким образом, пользоваться этим могут почти все. Мы можем узнать о поведении людей больше, чем когда-либо, а поведение тесно связано с решением проблем.

Как и у любой технологии, у «больших данных» есть обратная сторона. Сбор больших объемов данных может улучшить жизнь людей, но может быть использован и для управ-

ления ими, манипулирования и эксплуатации. Большинство организаций использует информацию с благими намерениями. Но не все разделяют такие намерения. Усиливается обеспокоенность по поводу влияния данных на людей, и многие правительства осознают необходимость в регулировании данных.

Для реализации преимуществ «больших данных» в развивающихся странах необходима революция как в использовании технологии, так и в методах ее применения. Только значительное изменение на обоих фронтах позволит «большим данным» в полном объеме способствовать

**Анализ данных
можно использовать
в качестве
локомотива
экономического
роста
в развивающихся
странах**



Фермер, выращивающий картофель, использует смартфон, Муншигондже, Бангладеш

развитию. Правительствам, наряду с неправительственными организациями (НПО) и поддерживающими их донорами и фондами, принадлежит решающая роль.

«Большие данные» быстро становятся задачей по постановке новых и более точных вопросов. Всякий раз, когда вы пользуетесь телефоном или что-то размещаете на платформе социальных сетей, вы создаете данные, которые накапливаются в море цифровой информации. Аналитики могут собирать эти данные, и вместо вопроса *что* произошло, могут теперь поставить более важный вопрос — *почему*. Чем объясняются поступки людей? Почему какие-то конкретные вводные данные дают конкретный результат? Ответы на эти вопросы могут на много пролить свет.

Например, пытаясь понять, чем объясняется большее число смертей на душу населения от болезней в конкретном районе, аналитик выясняет, что число пациентов, приходящих на прием к врачам в этом районе, уменьшается в определенное время года. Перекрестная проверка с другими данными указывает на корреляцию между посещениями врачей и снижением уровня воды в это время года. При низком уровне воды враги из соседнего района могут пройти на территорию, на которой расположена районная больница. В страхе за свою жизнь люди, даже заболев, избегают эту больницу. Новая больница или клиника в другом месте была бы разумным новым решением старой проблемы.

Практические решения

Программа HP Earth Insights, разработанная совместно организацией Conservation International и компанией Hewlett-Packard, значительно улучшает сбор, анализ и обмен данными о растительности, видах, запасах углерода и климате. Раньше ученые неделями изучали разрозненные данные и огромное количество, казалось бы, несвязанных фотографий диких животных в 17 тропических лесах в 15 странах. Пользуясь программным обеспечением «больших данных» для интеграции различных наборов данных, расчетов и моделирования, теперь они могут определить тенденции в считанные минуты и лучше понять, как потеря лесов и изменение климата ведут к утрате биологических видов. Затем Conservation International может предложить актуальные решения и определить, что работает лучше всего.

Или рассмотрим Kreditech, частную компанию, пользующуюся «большими данными» для оценки кредитных досье потенциальных заемщиков с помощью публично доступной информации, включая социальные сети. Это важно, поскольку у многих людей в развивающихся странах есть традиционное кредитное досье, что ограничивает их возможности по получению кредита вообще или по разумным ставкам. Альтернативные же методы Kreditech могут способствовать предоставлению необеспеченных кредитов заемщикам, которые ранее считались не соответствующими критериям. Так, например, если друг потенциального заемщика успешно получил ссуду, это считается положительным фактором при определении его собственной кредитоспособности. До сих пор коэффициент погашения ссуд Kreditech на некоторых рынках доходит до 93 процентов, указывая не только на легитимность ее критериев, но и на их потенциальное превосходство по сравнению с традиционными показателями. Возможность работать, не обращаясь к традиционным кредитным бюро, означает, что миллионы людей, не имеющих кредитного рейтинга, вскоре могут соответствовать критериям получения кредита.

К дополнительным примерам влияния «больших данных» на развитие относятся французская многонациональная телекоммуникационная корпорация Orange и шведский некоммерческий Фонд Flowminder, которые используют информацию, полученную из голосовых и текстовых данных мобильных телефонов, для мониторинга движения населения, чтобы знать, где лучше всего разместить центры лечения лихорадки Эбола.

«Большие данные» могут усовершенствовать анализ существующих проблем развития, но только при условии, что светлые головы возьмутся за решение этих вопросов и возьмут работу по устранению социально-экономических проблем. Здесь контекст данных заключается в разнице между открытием интересных фактов и фактов, на которые можно воздействовать в местных условиях. Недостаточно импортировать ученых-специалистов по данным из-за рубежа. Местные жители, знакомые с местными проблемами и культурой, способны лучше анализировать данные. Соответственно, международные доноры и правительства должны играть ключевую роль в поддержке подготовки местных аналитиков. Директивные органы, в свою очередь, должны быть готовы к экспериментам с творческими и смелыми новыми решениями проблем экономической политики, которые становятся возможными благодаря «большим данным».

Усиленная инъекция технологий

Многие люди задаются вопросом, могут ли «большие данные» заполнить пробелы, когда информация из традиционных государственных источников данных неполна или неточна. Мы уже многое можем узнать из данных, полученных в развивающихся странах благодаря телефонам, электросчетчикам и спутникам (например, о маршрутах поездок или хозяйственной деятельности). Но более развитая инфраструктура позволила бы сделать намного больше.

Дополнительные данные можно получить из усовершенствованной модификации мобильного телефона, который уже есть у большинства людей в странах с формирующимся рынком. По мере перехода растущего числа людей с мобильных телефонов на смартфоны будет расти объем актуальных данных. Это происходит потому, что смартфоны поощряют более интерактивное поведение с использованием приложений, обогащая «клад» для «раскопок» аналитиков. Поощрение принятия смартфонов на вооружение для решения проблем в конкретных секторах (таких как здравоохранение и сельское хозяйство) и местах (таких как города определенного размера) будет означать как рост количества, так и повышение качества информации в геометрической прогрессии.

Доноры и правительства не склонны к инвестированию в смартфоны для людей, которым они не по карману, поскольку они могут считаться предметом роскоши, как телевизор. Но ввиду ценности для общества получаемых данных, правительствам следует подумать о субсидировании стоимости смартфонов или связанных с ними абонементов на данные для сбора информации о пользователях. Расходы на субсидии можно возместить за счет этической продажи анонимных данных тем, кому они нужны в коммерческих целях. В мире сбора данных смартфоны не являются роскошью, они являются «последней милей» жизненно необходимой общественной инфраструктуры.

С увеличением числа владельцев смартфонов потребуются дополнительные приложения для производства соответствующего местного содержания. Приложения, помогающие людям найти чистую воду недалеко от дома, предупреждают о вспышке холеры или лихорадки Эбола или помогают им поддерживать местную школу, в свою очередь предоставляют дополнительную информацию, чтобы послужить пищей для еще более совершенного анализа данных и ответить на вопросы, о постановке которых мы даже не задумывались. В действительности нам неизвестно в полной мере, чего можно достигнуть, пока не начнется использование таких приложений.

Важное значение будут иметь приложения, разработанные и управляемые местными жителями, непосредственно знакомыми с потребностями своих соседей. Местные разработчики будут играть главную роль в этих усилиях, и их инициативы следует поддерживать. Расширение использования смартфонов создаст рынок для разработчиков приложений и расширит их ряды, а правительства, доноры и неправительственные организации

могут обеспечить обучение, место для работы и другие виды поддержки. Работу разработчиков программного обеспечения следует считать продолжением официальной статистической деятельности: их роль в сборе информации не менее важна.

Некоторые опасаются, что «большие данные» будут использоваться в странах с формирующимся рынком и развивающихся странах для укрепления авторитарных правительств, которые могут использовать эту информацию, чтобы сохранить власть. По меньшей мере, в результате использования собранных данных не по назначению может быть нарушена неприкосновенность личной жизни. Но ведь от применения любой технологии, а не от самой этой технологии, зависит хороша она или плоха.

Для укрепления доверия к сбору данных и обеспечения их надлежащего использования необходимы политика конфиденциальности и другие гарантии. Данные, собираемые в настоящее время, находятся у операторов мобильной связи, правительств, поставщиков приложений, сетевых провайдеров социальных медиа. Предоставление этих данных в агрегированной, анонимной форме для анализа данных чрезвычайно важно, и правительства могут сформулировать политику, позволяющую раскрытие этой информации (даже для частных поставщиков) в случае обоснованной необходимости.

Максимизация применения противомоскитных сеток

Некоторые утверждают, что инструменты «больших данных» и необходимые для них инвестиции не столь важны для развивающихся стран как, скажем, противомоскитные сетки. А что, если «большие данные» позволят оптимизировать сроки распределения противомоскитных сеток? На самом деле они уже в состоянии это делать. Информационная система для принятия решений по борьбе с комарами «aWhere's» сочетает изображения, полученные из спутниковых данных, погодные модели и биологию комаров для выявления потенциальных атак

еще до появления этих насекомых, — в конечном итоге предоставляя поставщикам и пользователям противомоскитных сеток бесценную информацию о том, где лучше всего использовать их запасы.

Этот пример показывает, как оптимизация использования данных может улучшить работу в области развития посредством обеспечения рационального использования ограниченных ресурсов в различных областях. «Большие данные» не сводятся к большей осведомленности, более глубокому пониманию или более точным прогнозам; речь идет о повышении отдачи при меньших затратах.

Для использования анализа данных в целях стимулирования роста в развивающихся странах необходимы следующие простые шаги: инвестиции в нужную технологию и образование, поддержка предпринимателей и сдвиг в менталитете, а также проведение политики конфиденциальности и открытого характера данных. Эту работу целесообразнее всего проводить одновременно, поскольку отставание в любой области задержит общее продвижение.

«Большие данные» могут и будут иметь значение в развивающихся странах, но как быстро и каким образом это произойдет, зависит от наших сегодняшних действий. Для реализации полного потенциала «больших данных», мы должны задействовать новые источники информации и внедрить новый подход к их использованию. Только тогда мы сможем ускорить развитие большего числа развивающихся стран с помощью этого нового мощного инструмента. ■

Кристофер У. Сурдак — автор книги «Data Crush: How the Information Tidal Wave Is Driving New Business Opportunities» и работает в сфере открытий и анализа в электронике, а Сара Агарвал — Директор по международным финансовым организациям в компании Hewlett-Packard.



PROGRAM IN ECONOMIC POLICY MANAGEMENT (PEPM)

Confront global economic challenges with the world's leading economists, policymakers, and expert practitioners, including Jagdish Bhagwati, Guillermo Calvo, Robert Mundell, Arvind Panagariya, and many others.

A 14-month mid-career Master of Public Administration focusing on:

- rigorous graduate training in micro- and macroeconomics
- emphasis on the policy issues faced by developing economies
- option to focus on Economic Policy Management or International Energy Management
- tailored seminar series on inflation targeting, international finance, and financial crises
- three-month capstone internship at the World Bank, IMF, or other public or private sector institution

The 2015–2016 program begins in July of 2015. Applications are due by January 5, 2015.

pepm@columbia.edu | 212-854-6982; 212-854-5935 (fax) | www.sipa.columbia.edu/academics/degree_programs/pepm
To learn more about SIPA, please visit: www.sipa.columbia.edu