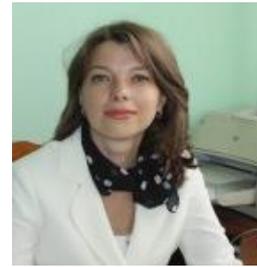


ЭТИКА БОЛЬШИХ ДАННЫХ



Д.А. Пархоменко

Старший преподаватель кафедры инженерной психологии и эргономики БГУИР, магистр техники и технологии



В.В. Шаталова

Заместитель декана по учебно-методической работе факультета компьютерного проектирования БГУИР, доцент кафедры проектирования информационно-компьютерных систем БГУИР, кандидат технических наук, доцент

*Белорусский государственный университет информатики и радиоэлектроники, Республика Беларусь
E-mail: parkhomenko@bsuir.by, shatalova@bsuir.by*

Анализ больших данных сегодня является не только трендом, но и необходимостью, которая может дать неоспоримое конкурентное преимущество. Технологии работы с большими данными могут принести огромную пользу как отдельным лицам, так и организациям, позволяя персонализировать обслуживание, обнаруживать мошенничество и злоупотребления, эффективно использовать ресурсы, предотвращать аварии или сбои в работе той или иной системы.

Одновременно с развитием технологии анализа больших данных встает вопрос об этике такой аналитики. По своей сути данная технология является этико-агностической, но это подталкивает к преодолению различных ограничений: наличия широкого диапазона данных из многих источников; возможность недорогостоящей корреляции этих данных для понимания более широкой картины; точность, с которой человек может быть идентифицирован и подвергнут таргетированию; умение определять местонахождение кого-либо для контекстного анализа и наблюдения; применение этого нового понимания для широкого круга решений и действий; работа в режиме реального времени.

Сегодня законы и нормативные документы не успевают отвечать тому, что порождают новые цифровые технологии. Конфиденциальность личных данных под угрозой. Также пользователям не хватает знаний и понимания таких технологий, чтобы защитить себя. Недавние достижения в аналитике больших данных расширили разрыв между тем, что возможно и что разрешено законом, изменили баланс сил между отдельными лицами и агрегаторами данных. Сегодня появляются новые возможности, связанные с социальными катастрофами и непредвиденными последствиями. Именно в этом промежутке поднимаются этические вопросы вокруг того, что приемлемо. Как организация применяет аналитику больших данных для повышения эффективности своей работы, понимает что использование этой технологии является этичным? Кто решает, как и когда правильно пользоваться большими данными? Согласно прогнозам аналитического агентства «Gartner», к 2018 году около половины всех нарушений деловой этики будет обусловлена неправильным использованием имеющихся средств анализа больших данных.

Примеры хорошей и плохой практики появляются в отрасли, и со временем они будут определять правила и законодательство.

История сети супермаркетов «Target», о том что алгоритм определил беременность девочки и стал присылать ей купоны на товары для беременных, еще до того как семья девочки узнала о ее беременности, часто звучит как пример возможностей больших данных и машинного обучения. Однако эта история подняла волну протеста, и «Target» потеряла лояльность своих клиентов.

Проблемы, которыми занимается цифровая этика, можно объединить в несколько групп:

- проблема «Выбора» – когда вам нужно выбрать один из нескольких вариантов, каждый из которых является приемлемым для вас и не касается вас лично;
- проблема «Собственника» – когда вы должны сделать личный выбор – вы хотите заработать много денег или сделать вашего клиента счастливым;
- проблема «Времени» – когда вы должны выбрать между тактическим и стратегическим решением.

Профессор права Вашингтонского университета Нил Ричардс отмечает: «этика диктует принципы, которыми руководствуются законодатели». Принятие законов, касающихся цифровой сферы сильно запаздывает. Сегодня закон практически не регламентирует манипуляции с большими данными, хотя уже возникают моменты, которые люди воспринимают с возмущением.

Для того, чтобы компаниям избежать потери лояльности клиентов в новых рыночных условиях использования аналитики больших данных, компаниям необходимо вести себя прозрачно и обсуждать вопросы использования данных непосредственно с потребителями.