

# Большие данные (BigData) в естественно-научных исследованиях.

*А.П.Крюков<sup>1</sup> (НИИЯФ МГУ, Москва)*

## Абстракт

В докладе ставится проблема обработки и анализа с больших данных в естественно-научных исследованиях. Примерами таких областей могут служить физика, геология, биология и другие.

Особенностью обработки и анализа больших данных в естественных науках является то, что источники больших данных в этих областях — это экспериментальные установки, данные с которых представляют собой неструктурированные объекты типа изображений (возможно в абстрактных пространствах) и/или временных последовательностей. Для обработки и анализа такой информации необходимо решить несколько задач: хранение неструктурированной информации, обработка запросов на выборку по имеющейся метаинформации, организация обработки данных с целью извлечения скрытой структуры данных, реорганизация данных с учетом результата их анализа.

Успешное решение указанных задач возможно только обеспечив конечного пользователя удобным инструментарием, который обеспечивает поддержку всего цикла работы с большими данными.

Отличительной особенностью современного эксперимента является необходимость использования множества различных источников больших данных для их совокупной обработки и анализа. Это, в свою очередь, требует привлечение технологий распределенного хранения и обработки данных, а также использование суперкомпьютерных вычислений. В докладе также показано ключевое значение реорганизации данных с использованием результатов анализа в качестве метаинформации.

В качестве основного примера методов решения поставленных задач рассматривается проблема обработки и анализа экспериментальных данных с Большого адронного коллайдера.

---

<sup>1</sup> kryukov@theory.sinp.msu.ru