

Встреча на физфаке МГУ
13 апреля 2016

Федор Ратников,
Яндекс

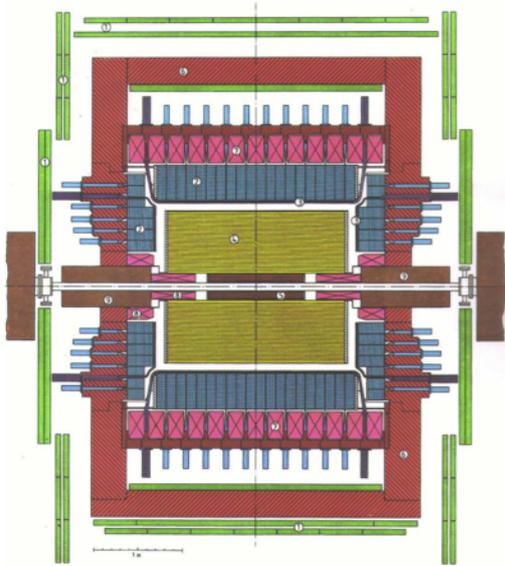
Яндекс и Физические Основы Мироздания

План

- ◇ о физике
- ◇ об Яндексе
- ◇ о карьере

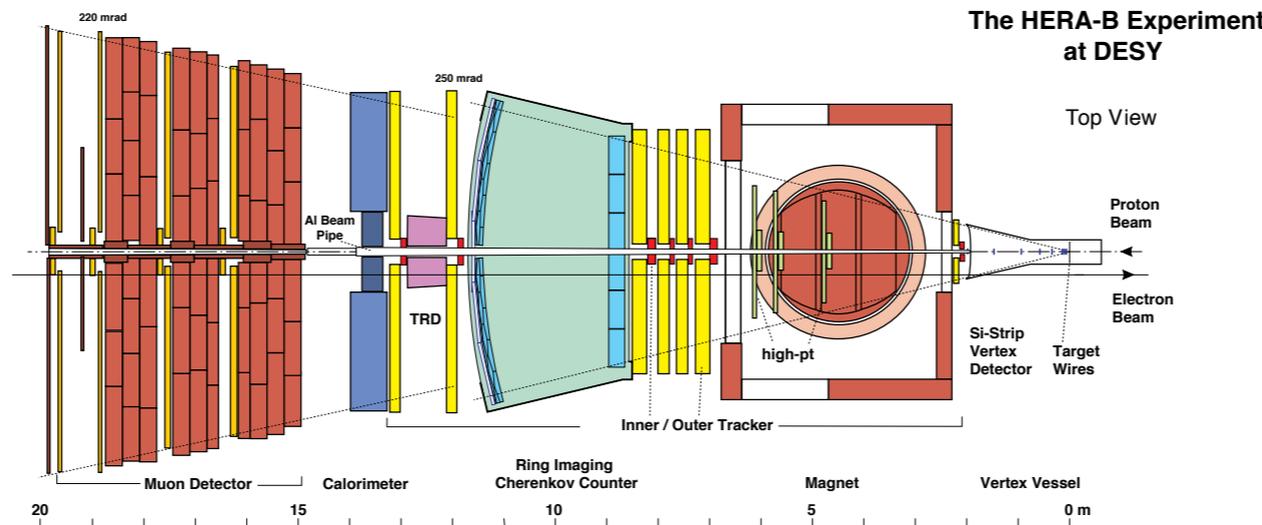
Мои 30 лет в Физике Высоких Энергий

ARGUS@DESY
Habmurg, DE
1989-1992

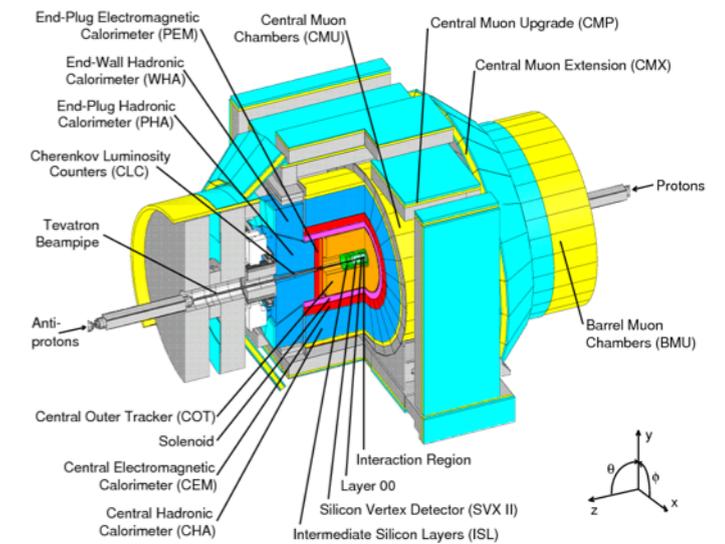


- 1 Muon chambers
- 2 Shower counters
- 3 Time of flight counters
- 4 Drift chamber
- 5 Vertex chamber
- 6 Iron yoke
- 7 Solenoid coils
- 8 Compensation coils
- 9 Mini beta quadrupole

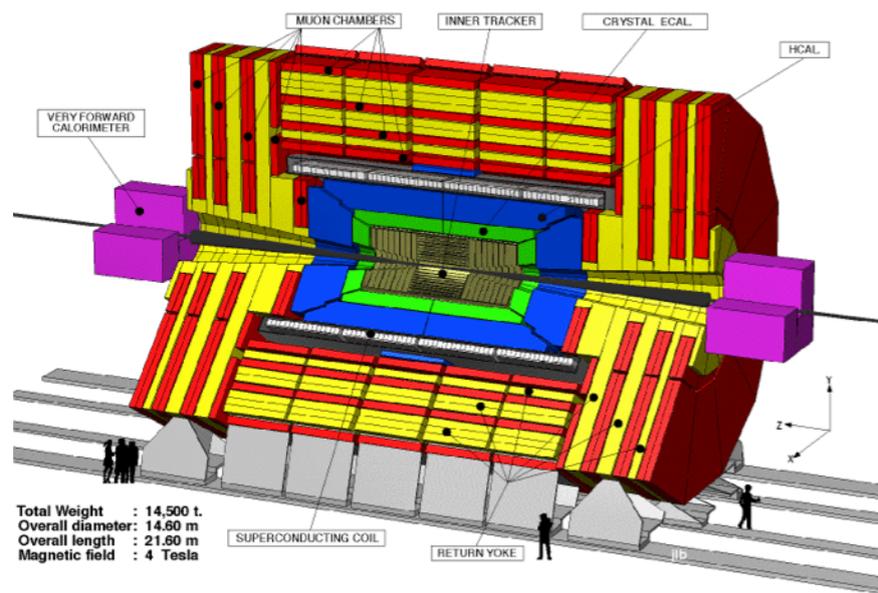
HERA-B@DESY
Habmurg, DE
1994-1999



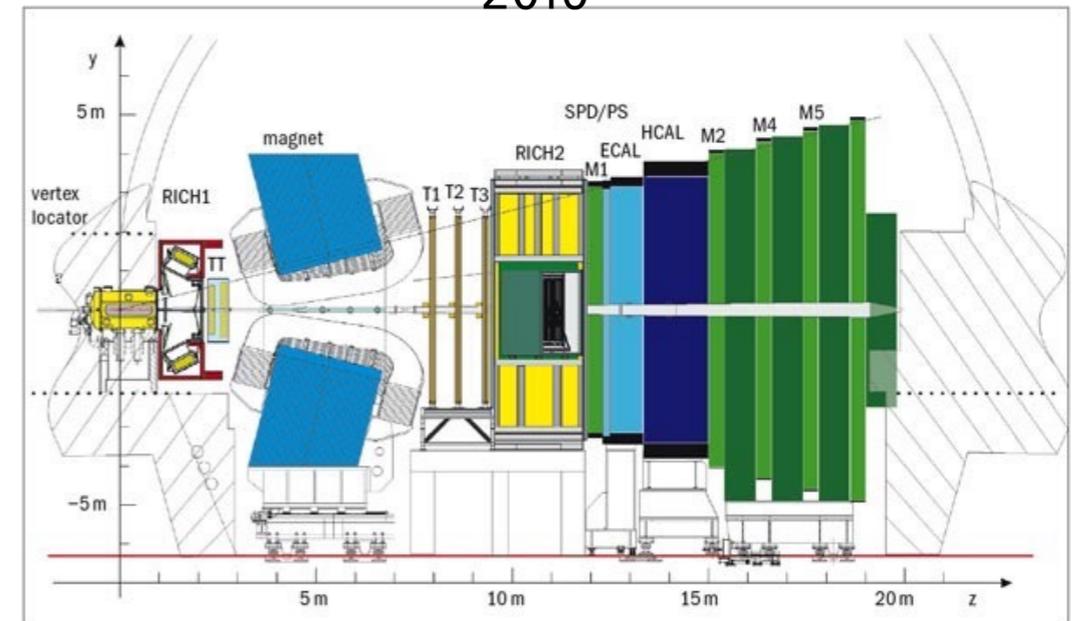
CDF@Fermilab
Batavia, USA
1999-2005



CMS@CERN
Geneva, CH
2005-2015



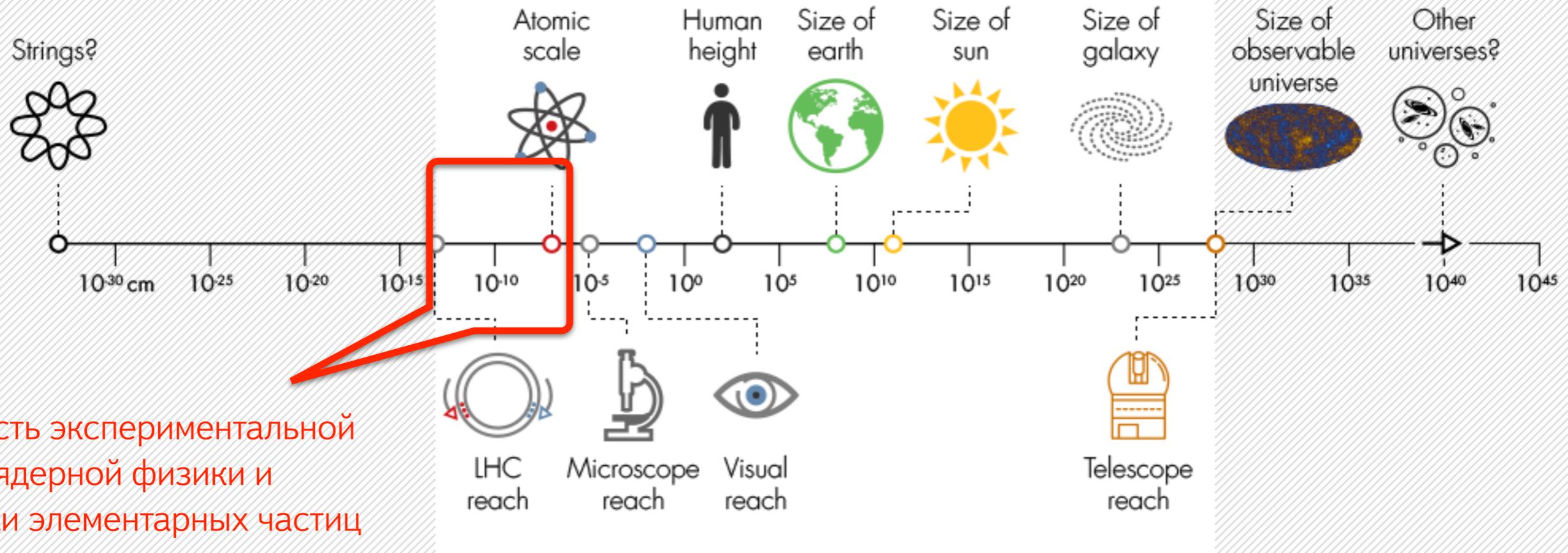
LHCb@CERN
Geneva, CH
2016-



Шкала Размеров

The Ends of Evidence

Humans can probe the universe over a vast range of scales (white area), but many modern physics theories involve scales outside of this range (grey).



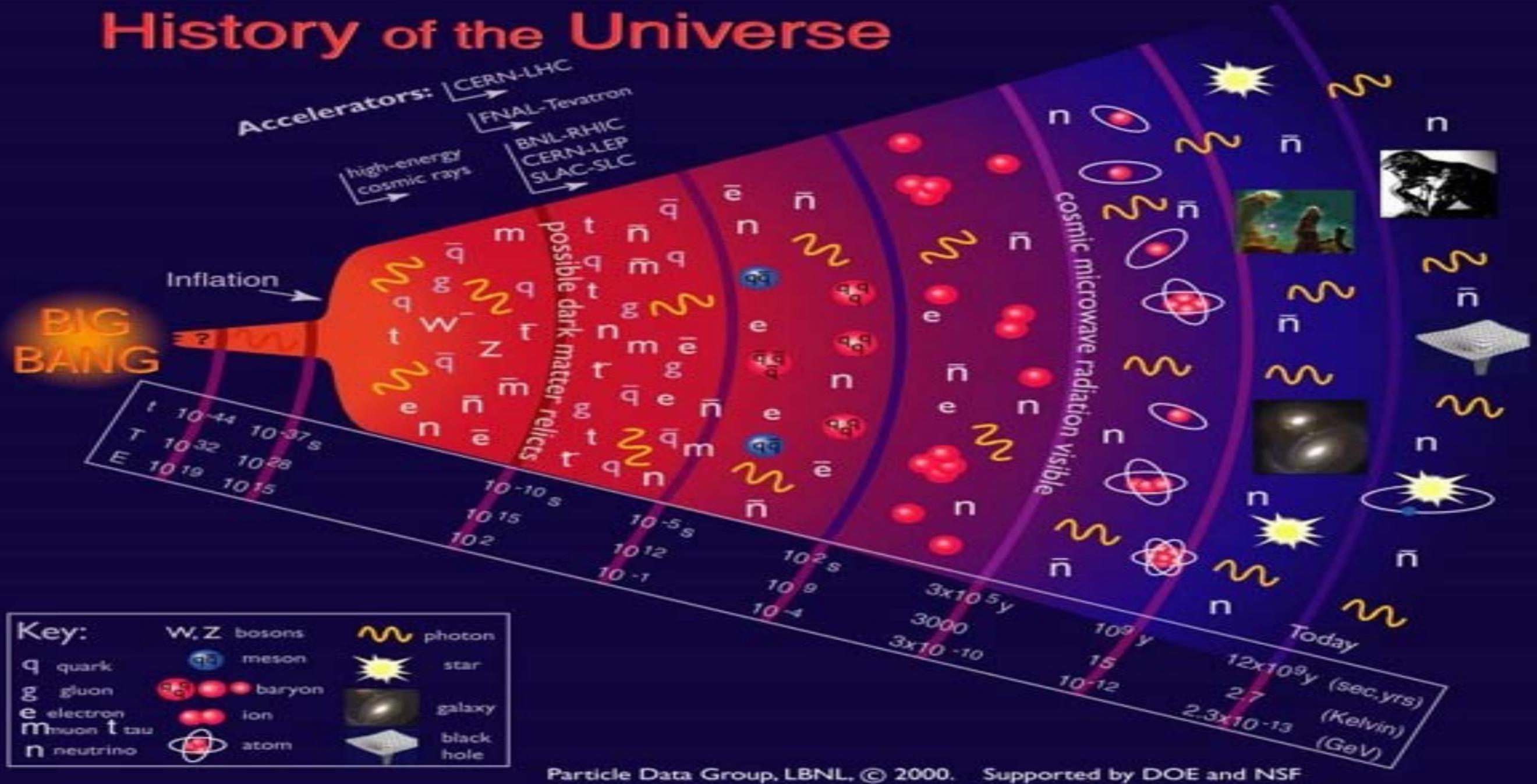
Область экспериментальной ядерной физики и физики элементарных частиц

◇ Как исследовать структуры малых размеров?

◇ $\Delta x \cdot \Delta p_x \geq h/2$

◇ использовать пробные частицы больших энергий!

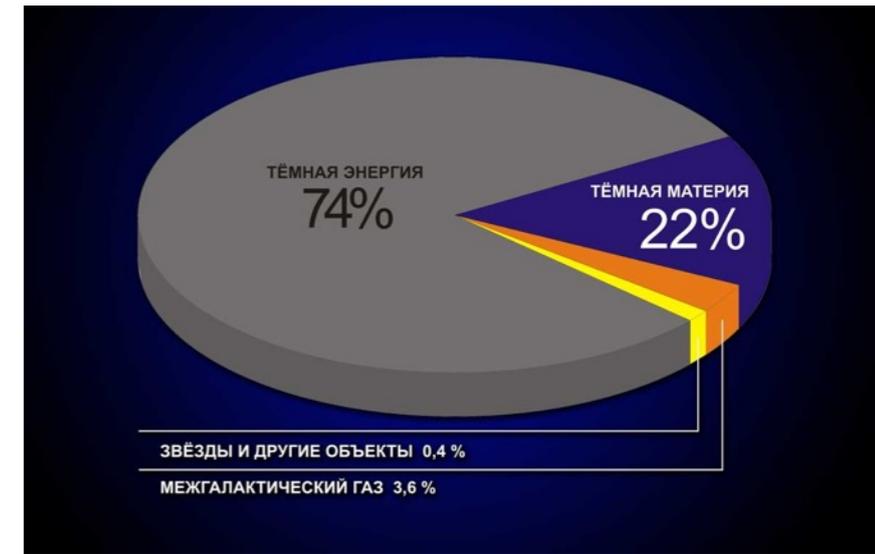
История Вселенной



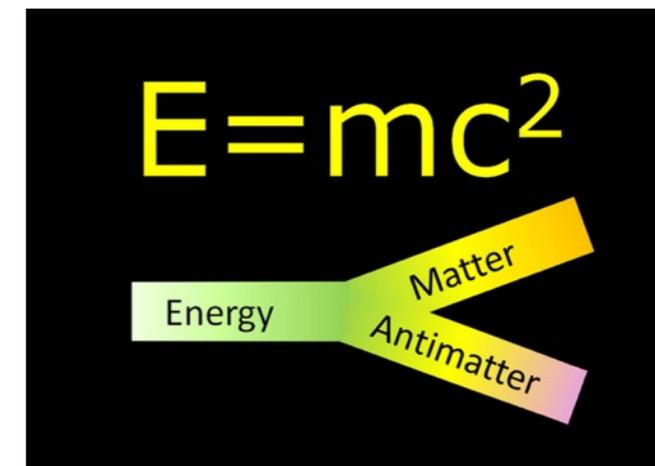
- ◇ Большие энергии частиц соответствуют состоянию более ранней Вселенной
- ◇ исследуя сегодня столкновения при больших энергиях мы заглядываем во Вселенную возраста $\sim 10^{-10}$ с (!)

Фундаментальные Проблемы Физики

- ◇ Темная материя и энергия
 - ◇ почему динамика Вселенной не соответствует распределению видимого вещества?
- ◇ Точная подгонка параметров
 - ◇ почему массы частиц такие, какие они есть (eV, MeV, GeV)?
- ◇ Сильная асимметрия между материей и антиматерией
 - ◇ куда подевались антимирры?
- ◇ Масса частиц
 - ◇ откуда берется масса частиц?



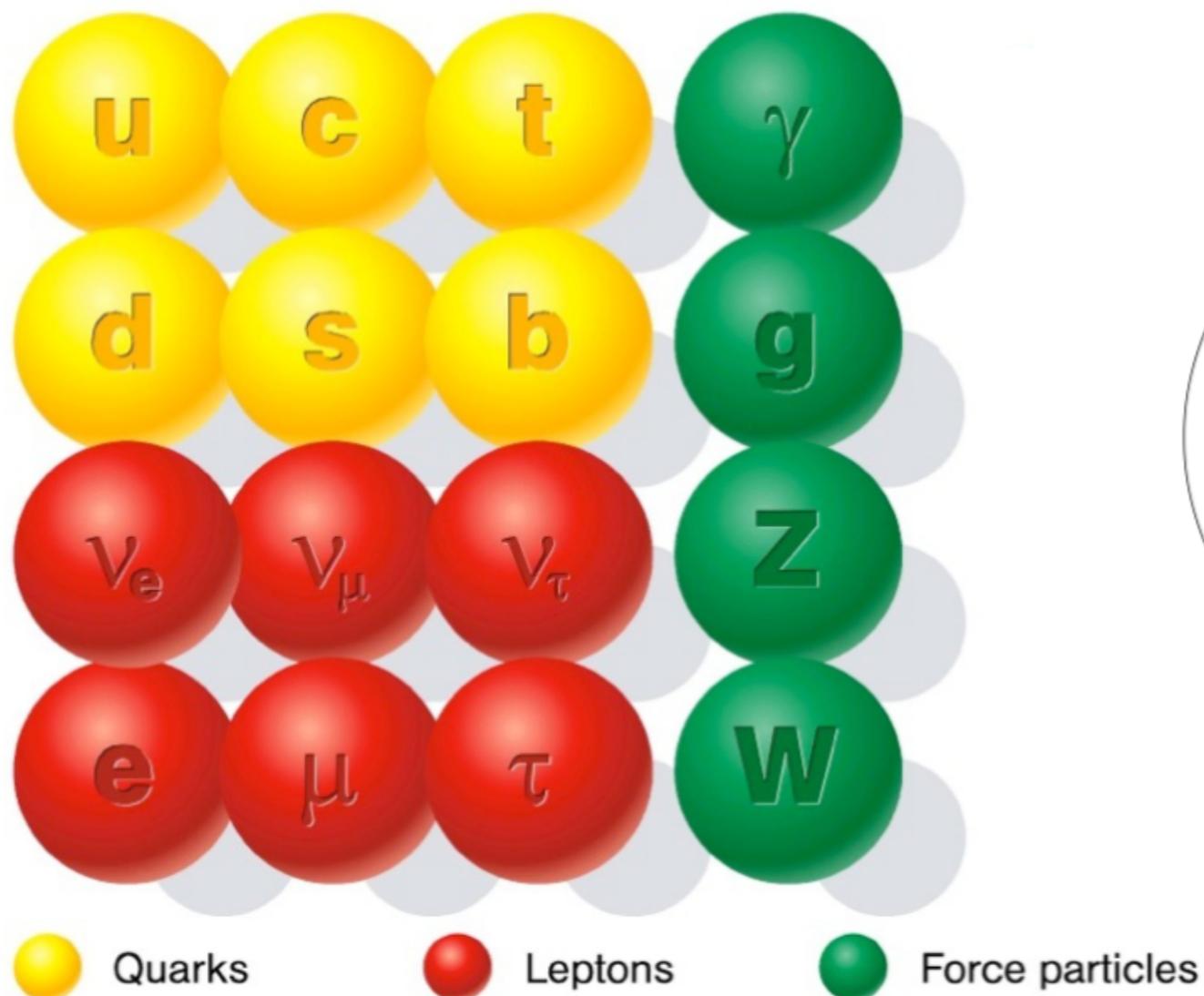
$$\sqrt{\frac{\hbar c}{8\pi G}} \approx 4,340 \text{ мкг} = 2,43 \cdot 10^{18} \text{ ГэВ}/c^2$$



SOLVED!

Теория: Стандартная Модель

Standard particles



- ◇ SM - математическое описание материи как набора частиц и взаимодействий между ними

взаимодействие = обмен “частицами взаимодействия”

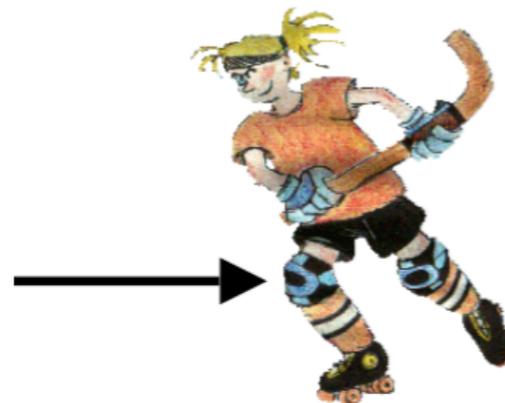


◇ игрок 1 бьет по мячу

◇ игрок 2 ловит мяч

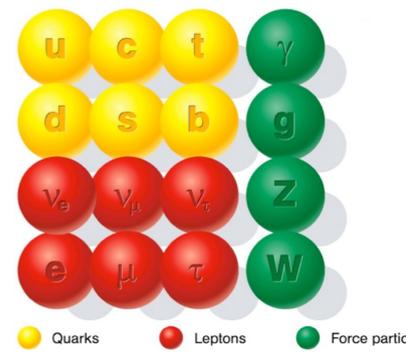


отталкивающая
сила

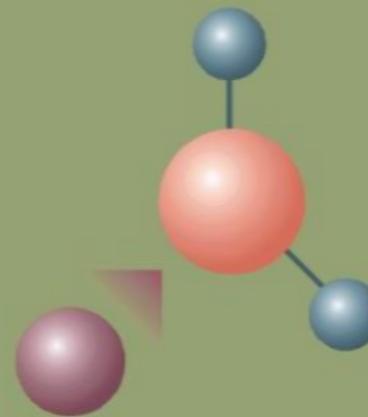
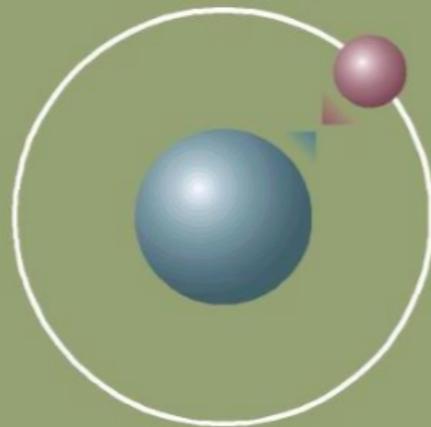
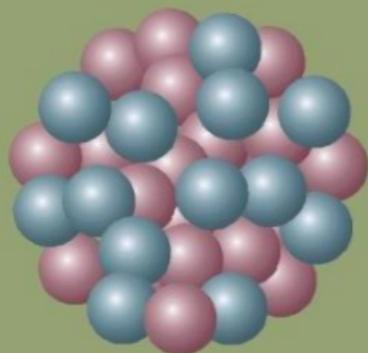


◇ оба игрока
разъезжаются
друг от друга

Взаимодействия



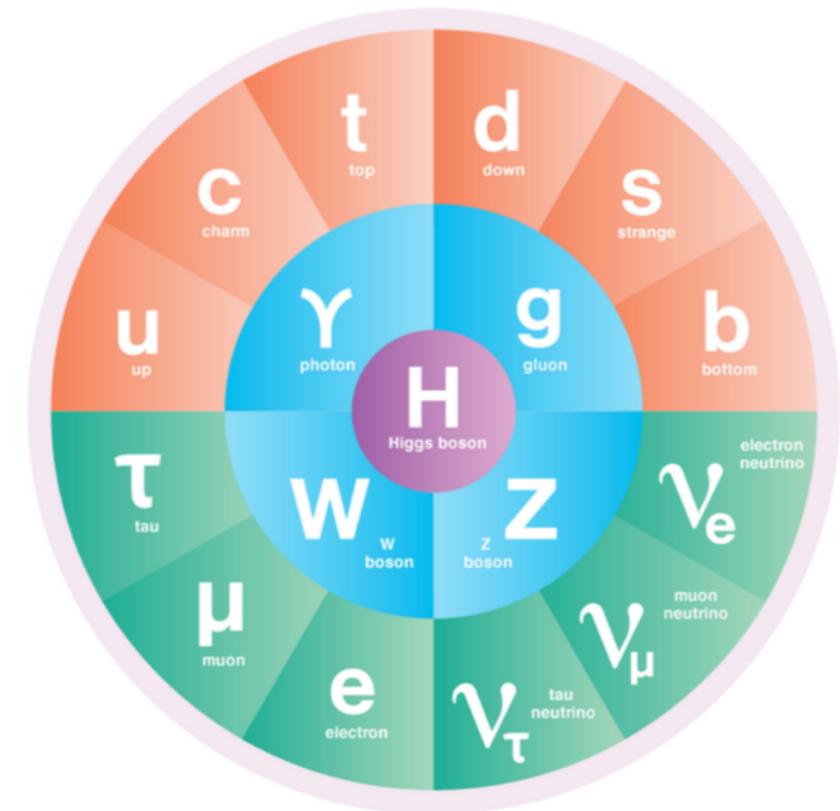
Gluons	Photon	W and Z boson	Graviton
Carriers of the:			
Strong force	Electromagnetic force	Weak force	Gravitational force
Affecting:			
Quarks, gluons	Quarks, charged leptons and W bosons	Quarks and leptons	All particles
Responsible for:			
Holding together the proton, the neutron and atomic nuclei	Chemistry, electricity and magnetism	Radioactivity, processes in the sun	Holding together the earth, the sun, the planetary system



Бозон Хиггса

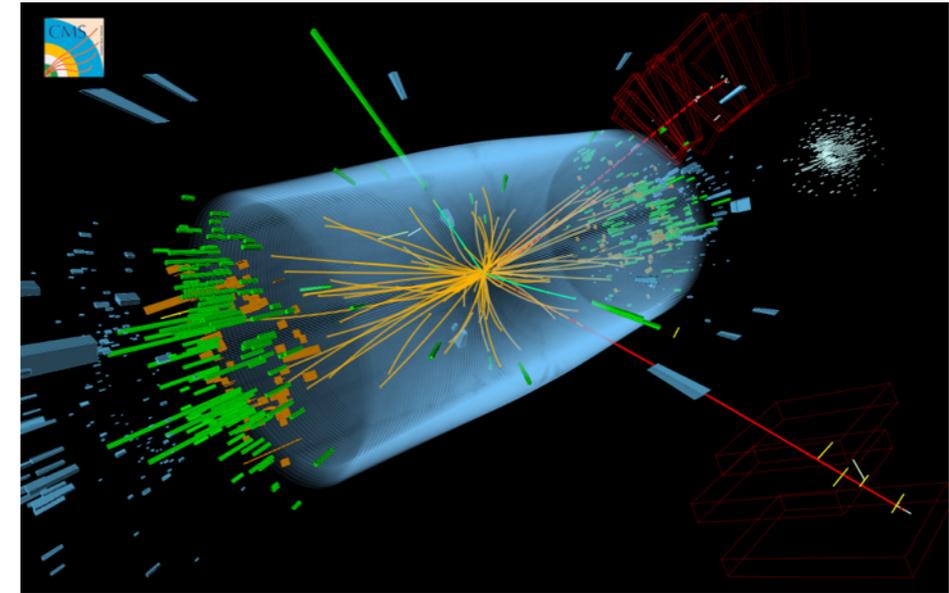


- ◇ Математически объясняет возникновение массы частиц
- ◇ Является одновременно частицей и полем, заполняющим всю Вселенную
- ◇ Взаимодействует с другими частицами - чем сильнее, тем частица тяжелее

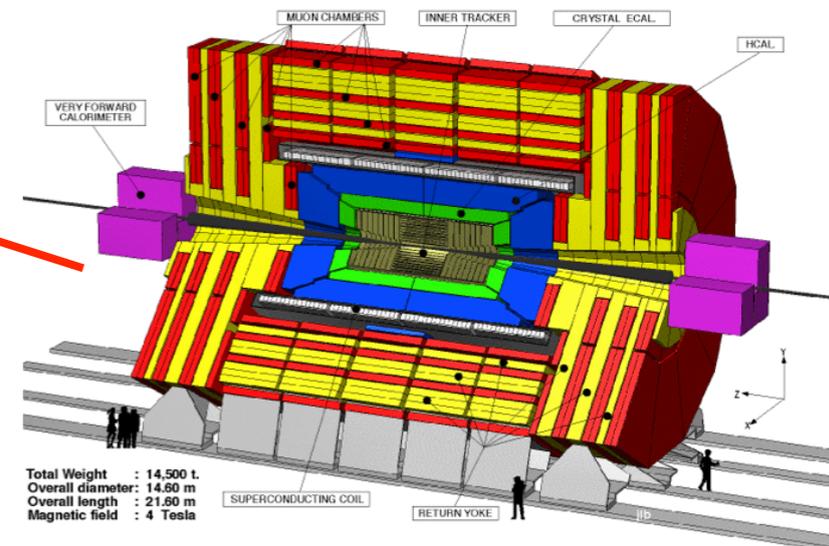
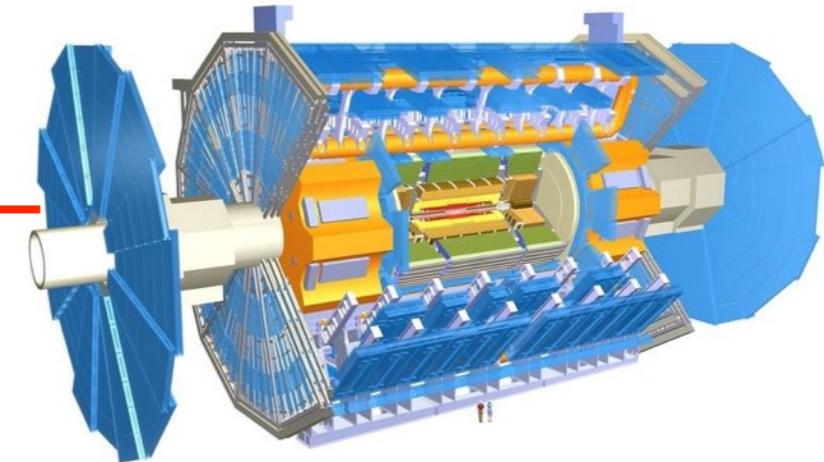
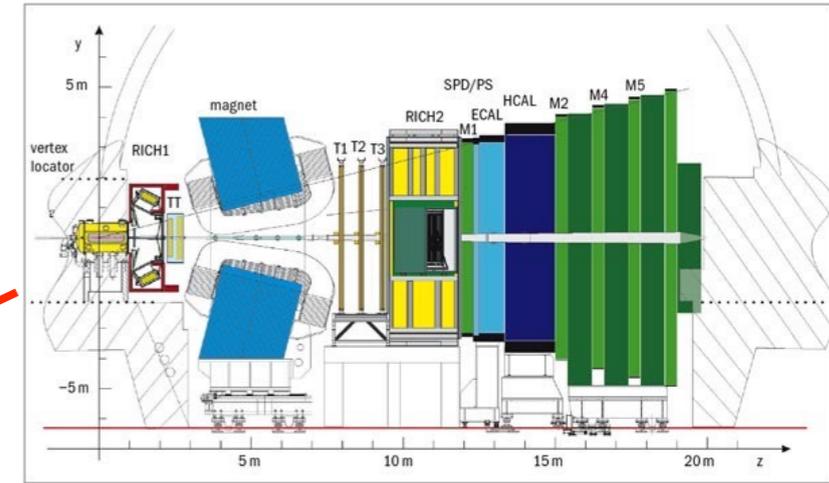
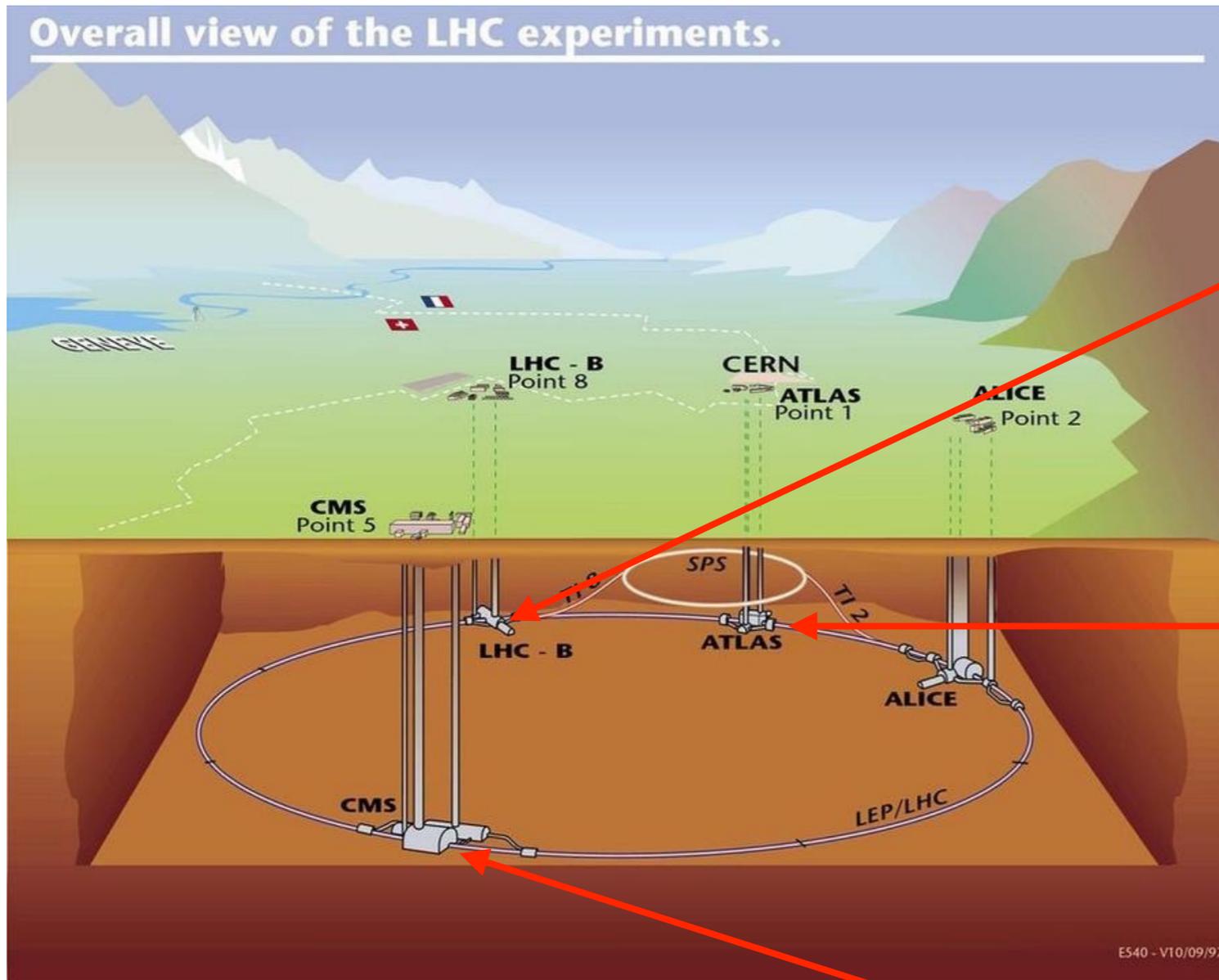


Откуда мы их берем?

- ◇ Буквально выколачиваем из ничего
 - ◇ создаем много энергии в малой области пространства
- ◇ После рождения они мгновенно распадаются на обычные частицы
- ◇ Мы не можем видеть сам бозон Хиггса, мы можем зарегистрировать продукты его распада
- ◇ Нам нужны:
 - ◇ ускоритель частиц, чтобы сконцентрировать энергию
 - ◇ детектор частиц, чтобы зарегистрировать продукты распада



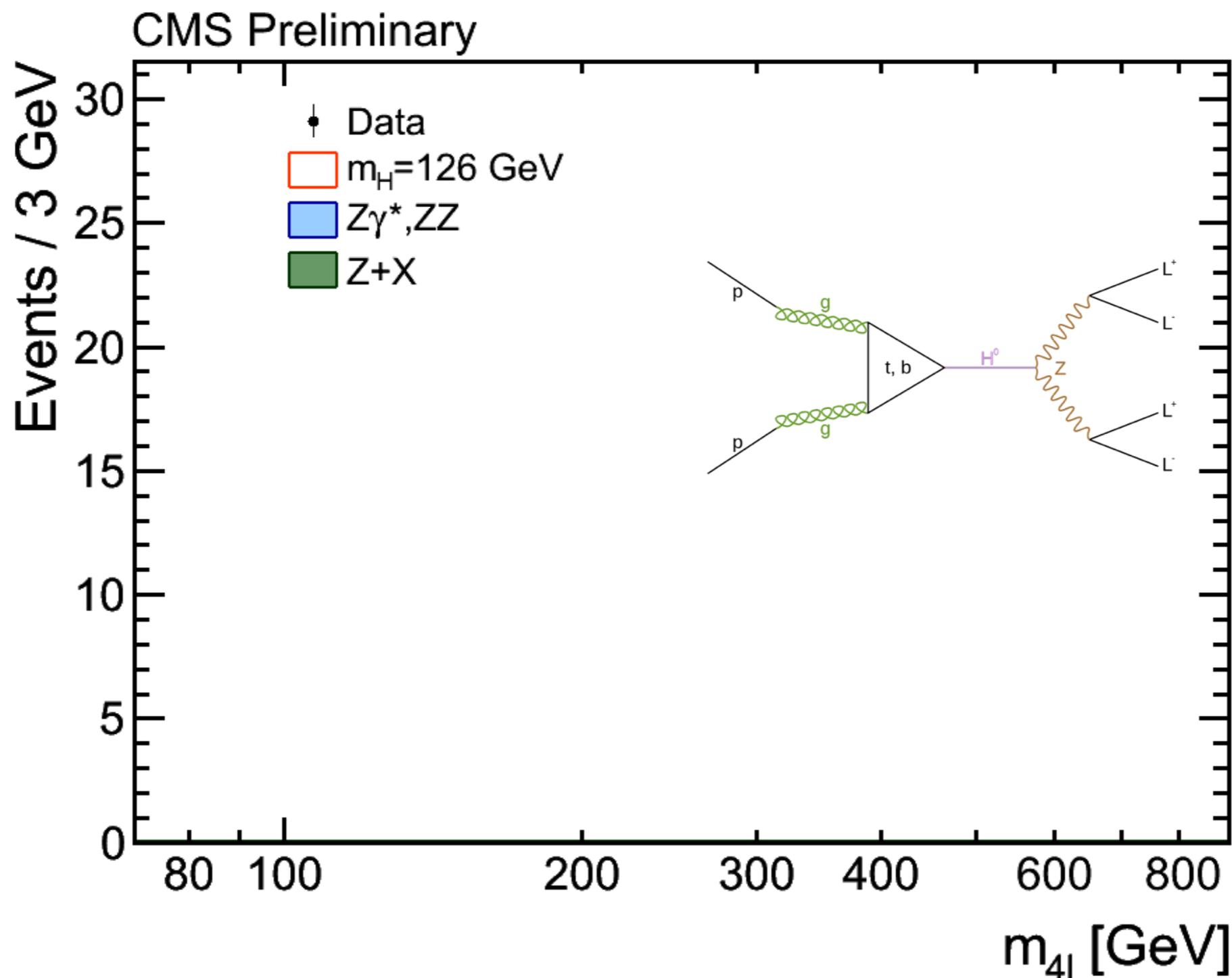
Большой Адронный Коллайдер (LHC)



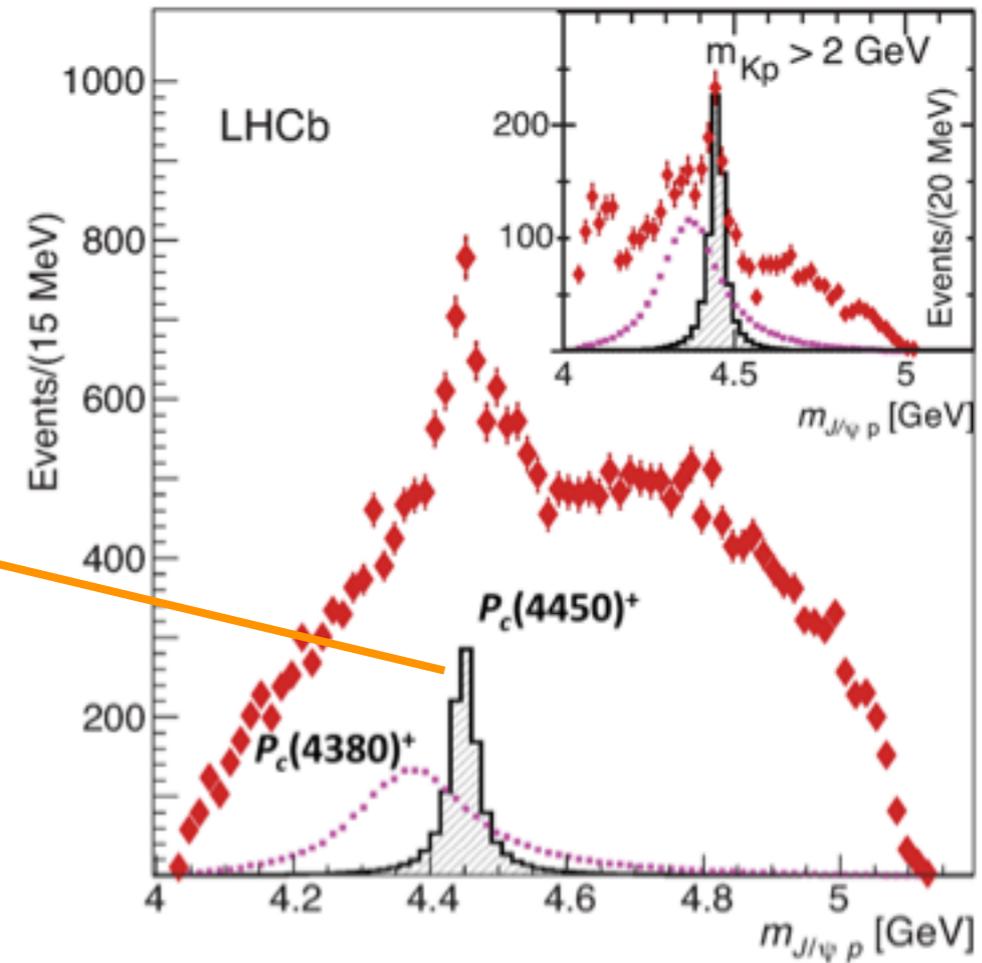
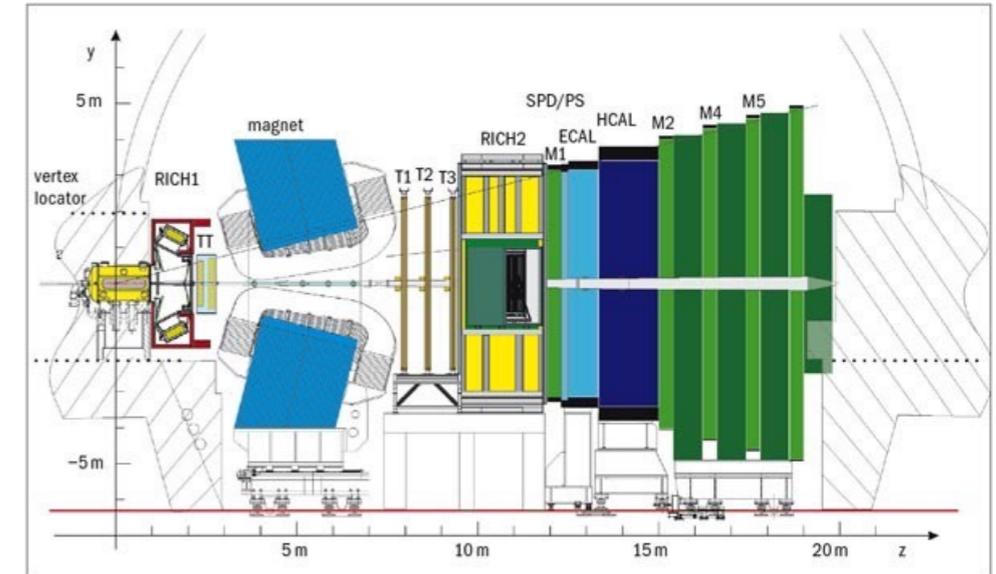
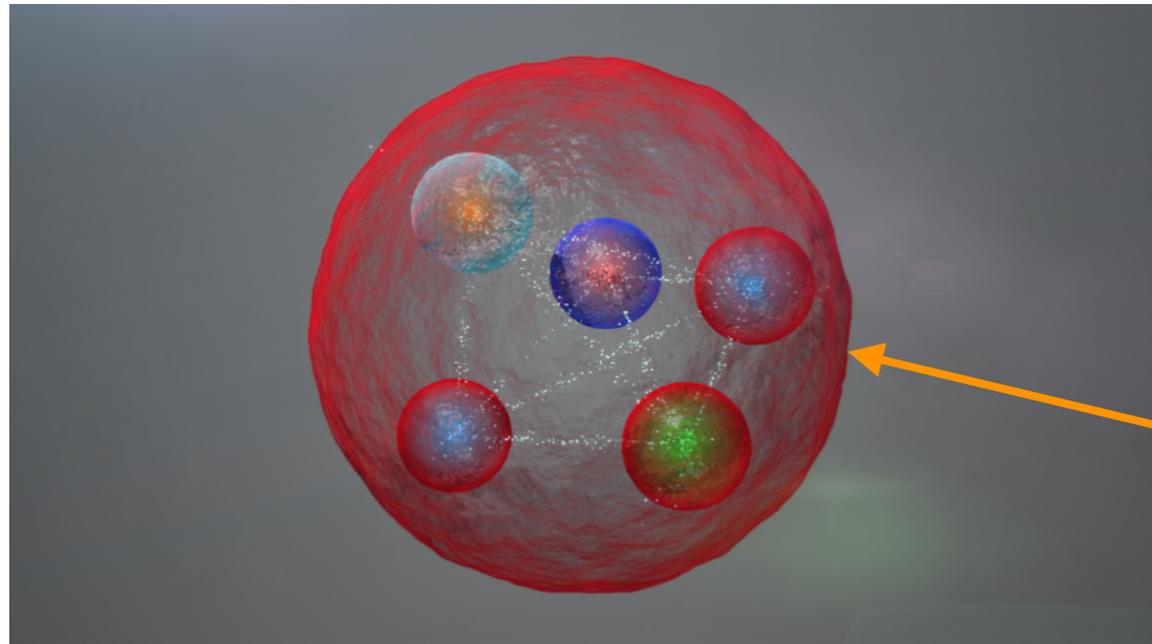
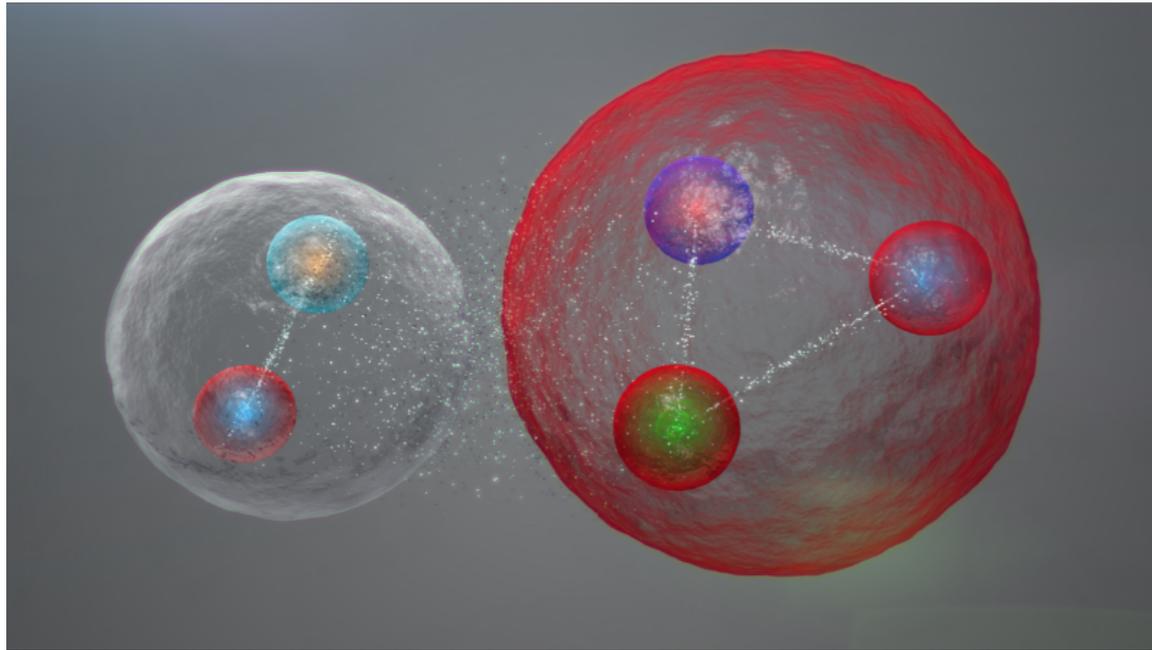
- ◇ 2 протонных пучка, скорость - 99.99999991% скорости света
- ◇ 4 точки пересечения пучков
- ◇ энергия в пучках ~ энергии Титаника на полном ходу

Открытие Бозона Хиггса в 2012

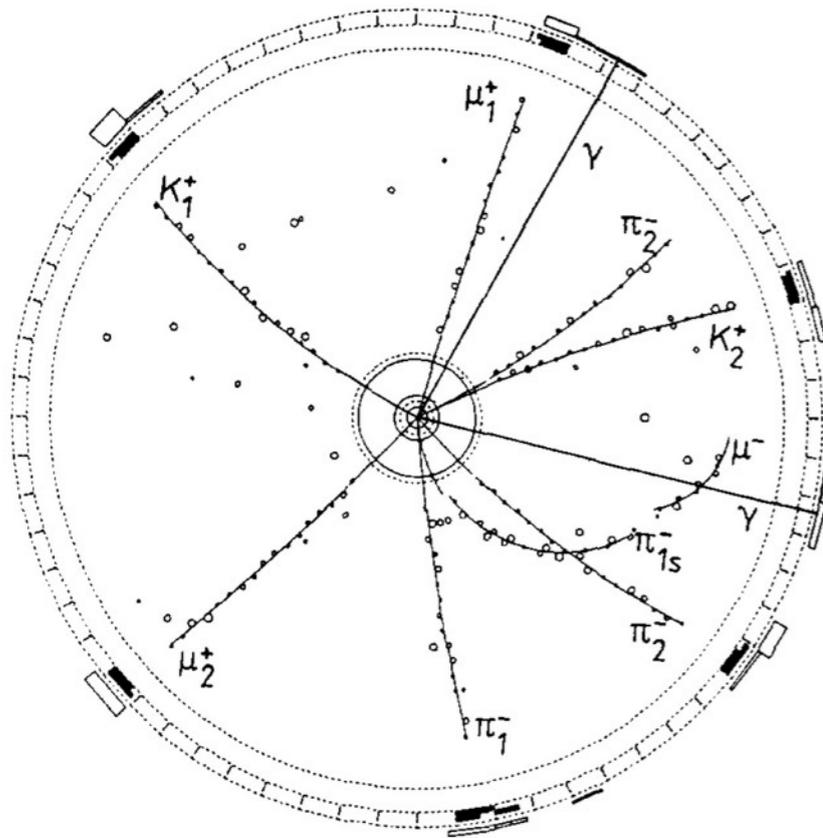
- ◇ Эксперимент = сравнение наблюдений с ожидаемым эффектом



Открытие Пентакварков на Детекторе LHCb

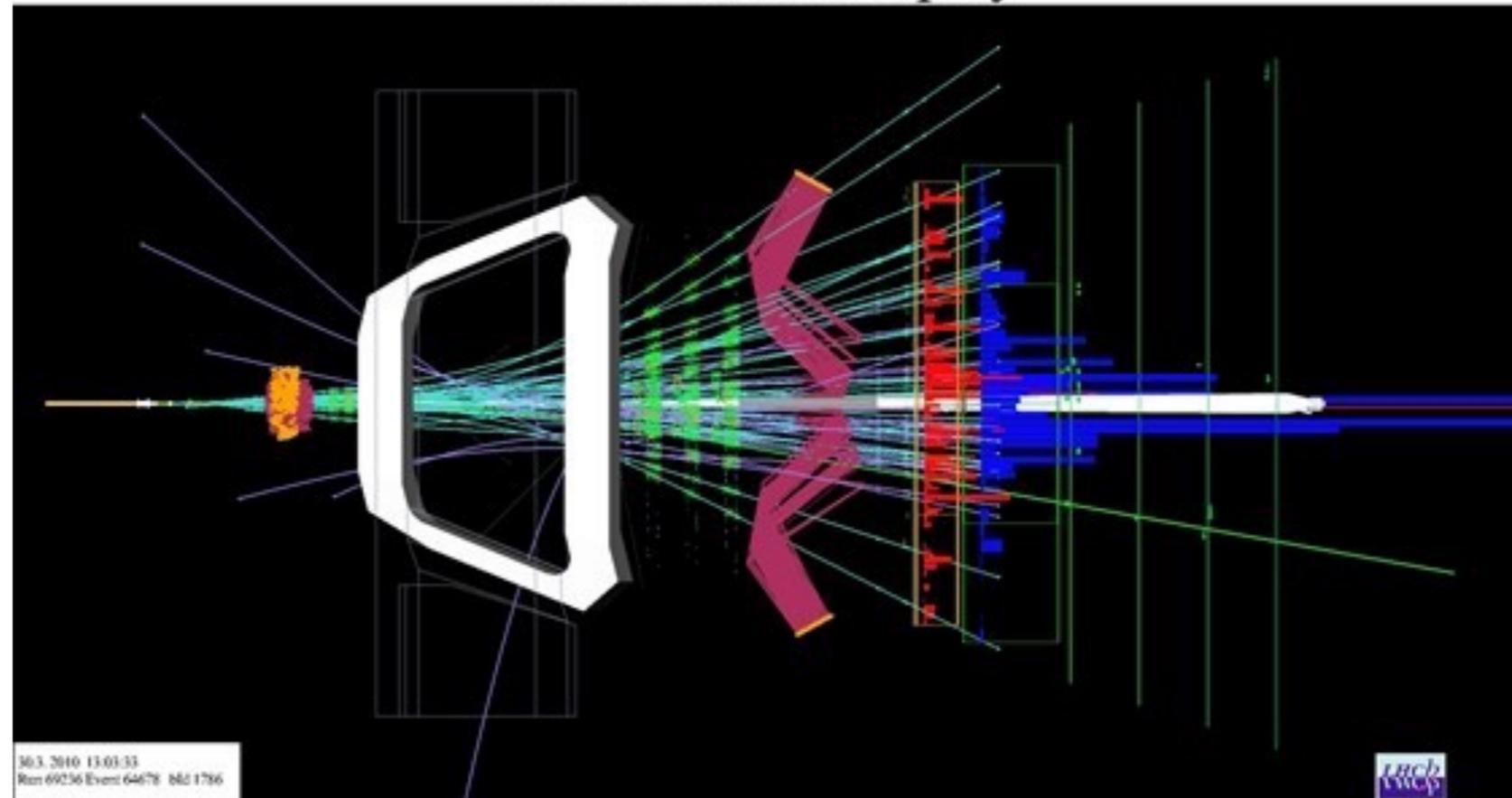


Реконструкция Информации в Детекторе



@ few Hz
ARGUS@DESY
1990s

LHCb Event Display



@ 40 MHz
LHCb@CERN
2010s

- ◇ Эксперименты накапливают десятки Петабайт физических данных
- ◇ Требуются быстрые, надежные, эффективные методы реконструкции параметров частиц и целых событий

Миссия Яндекса (Выдержка)

- ◇ Яндекс — технологическая компания. В основе наших сервисов лежат сложные, уникальные, трудно воспроизводимые технологии. Именно они позволяют нам делать то, что еще некоторое время назад люди приняли бы за волшебство.
- ◇ Наука. Нам удалось собрать команду специалистов во многих областях науки — в математике, анализе данных, программировании, лингвистике и других дисциплинах. Вычислительные возможности и алгоритмы Яндекса используют и наши партнеры для проведения своих научных исследований — например, в области ядерных исследований и геологоразведки.

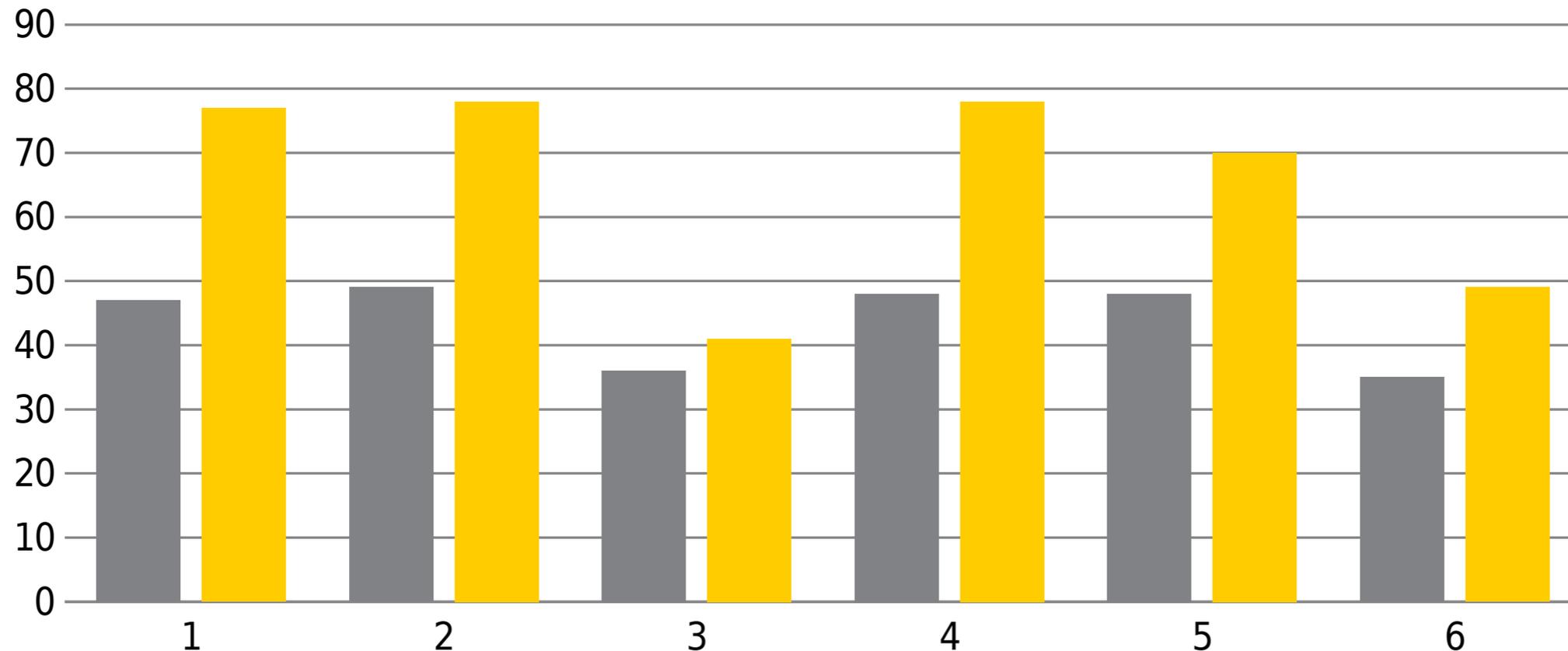
Сотрудничество ЦЕРНа и Яндекса

- ◇ Физике высоких энергий требуются эффективные методики извлечения уникальной информации из большого количества накопленных данных
- ◇ Современные IT компании нацелены в первую очередь на эффективное извлечение полезной информации из больших наборов данных
- ◇ Естественна заинтересованность физического сообщества в использовании наработок IT компаний для анализа физических данных
- ◇ Яндекс является безусловным лидером в отношении анализа больших данных, методов машинного обучения, искусственного интеллекта в России
- ◇ Яндекс заинтересован в продвижении и опробовании своих наработок в таких сложных областях, как анализ больших физических данных
- ◇ Результат - сотрудничество Яндекса и ЦЕРНа в обработке и анализе данных, прямое участие группы Яндекса в физических исследованиях коллаборации LHCb, включая полный доступ к имеющимся физическим данным для анализа

Истории Успеха

N-Body trigger Performance Comparison

(bars correspond to trigger efficiency in % for different decay modes)



■ Run-I (Before optimization) ■ MatrixNet

- ◇ Использование MatrixNet (поисковый движок Яндекса) позволило существенно улучшить эффективность триггерного отбора событий

Истории Успеха

- ◇ Данные хранятся на дисках (быстро, но дорого) и /или на лентах (дёшево, но медленно)
- ◇ С использованием машинного обучения удалось предсказать востребованность различных файлов и соответственно реорганизовать их на диске /ленте
- ◇ Результат: экономия 40% данных LHCb при ошибке предсказания ~1%

Специфика

- ◇ Современная экспериментальная физика высоких энергий означает работу на огромном детекторе в огромной коллаборации



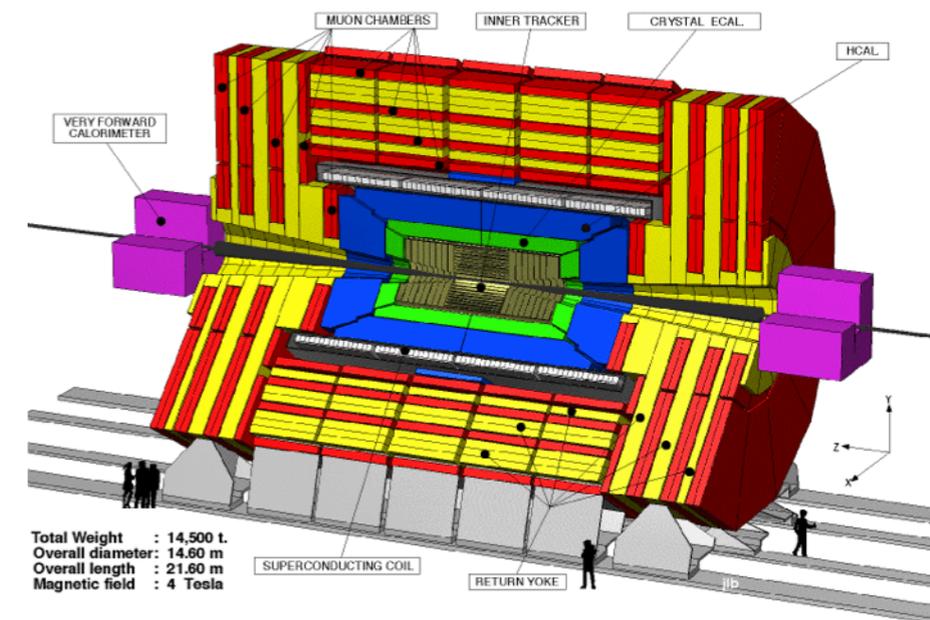
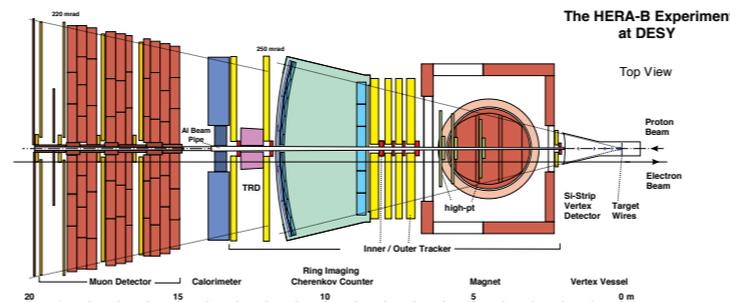
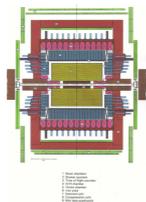
1990-e



2000-e



2010-e



Экспериментальная Физика Высоких Энергий

◇ Преимущества

- ◇ исследования основополагающих законов Мироздания
 - ◇ открытия наиболее базовых принципов
- ◇ работа в большой (100..1000 участников) коллаборации
 - ◇ постоянное общение и обмен идеями с многими выдающимися коллегами, выработка навыков решения новых, нетривиальных задач
 - ◇ навык работы в коллективах с горизонтальными связями, как больших, так и стартапов: организация взаимодействия, совместное планирование, обсуждение и претворении в жизнь идей
- ◇ использования средств накопления и анализа большого количества данных
 - ◇ Один из основополагающих навыков в современном пост-индустриальном обществе

Экспериментальная Физика Высоких Энергий

◇ Сложности

- ◇ исследования основополагающих законов Мироздания
 - ◇ требует изучения далеко продвинутых дисциплин теоретической, феноменологической и экспериментальной физики
- ◇ работа в большой (100..1000 участников) коллаборации
 - ◇ необходимость грамотной координации своей работы и интересов с работой и интересами множества коллег
 - ◇ навык работы в большом коллективе с горизонтальными связями может оказаться бесполезным в организациях с сильными вертикальными связями
- ◇ использования средств накопления и анализа большого количества данных
 - ◇ требует изучения продвинутых дисциплин прикладной и вычислительной математики, компьютеринга, машинного обучения и т. п.

Школа Анализа Данных Яндекса

- ◇ “В Школе, основанной Яндексом в сентябре 2007 года, преподают машинное обучение, компьютерное зрение, анализ текстов на естественном языке и другие направления современных компьютерных наук. Ежегодно в США поступают старшекурсники, выпускники и аспиранты МГУ, МФТИ, ВШЭ, ИТМО, СПбГУ, УрФУ, НГУ и других ведущих вузов. Два года студенты изучают предметы, которые обычно не входят в университетские программы, хотя пользуются огромным спросом в науке и разных отраслях, где уже применяются наукоёмкие информационные технологии.”
- ◇ Обучение фундаментальным основам и практическому использованию современных средств анализа данных и машинного обучения, применимых как к физике, так и другим практическим областям

- ◇ Приглашаем амбициозных студентов-физиков попробовать себя в условиях реальной научной работы

Студент Нашей Мечты

- ◇ Заинтересован в познании физического устройства окружающего мира, его базовых принципов
- ◇ Готов приложить серьезные усилия для изучения необходимых физических дисциплин, теорий, методик
- ◇ Любит и умеет осваивать и продвинутое компьютерные технологии использовать их в текущей работе
- ◇ Заинтересован в изучении передовых методик и средств анализа больших данных, машинного обучения
- ◇ Ставит целью пройти путь вплоть до защиты Ph.D. с дальнейшим построением академической карьеры, либо переходом в индустрию или стартап

Мы Обеспечиваем

- ◇ Темы работ, основанные на участии в исследованиях мирового класса
 - ◇ на уровне бакалавриата, магистратуры, аспирантуры
- ◇ Непосредственный доступ к данным международных коллабораций для анализа
- ◇ Изучение необходимого комплекса физических наук
 - ◇ на кафедре физики атомного ядра и теории столкновений физфака МГУ
- ◇ Изучение передовых методик анализа больших данных
 - ◇ в рамках общения с коллегами в группе Яндекса
 - ◇ в рамках Школы Анализа Данных Яндекса
 - ◇ ... верхний предел практически не ограничен
- ◇ Материальная помощь при условии успешной работы