

物理学との対話2
— 科学とAIの接点 —

「素粒子物理実験」とAIの接点

田中 純一
東京大学
素粒子物理国際研究センター

2019年6月5日
JSAI2019日本物理学会・人工知能学会連携企画セッション

ヒッグス発見と素粒子の標準理論

2012年7月4日



ヒッグス粒子発見か
新素粒子検出 年内に結論
日米欧2チーム

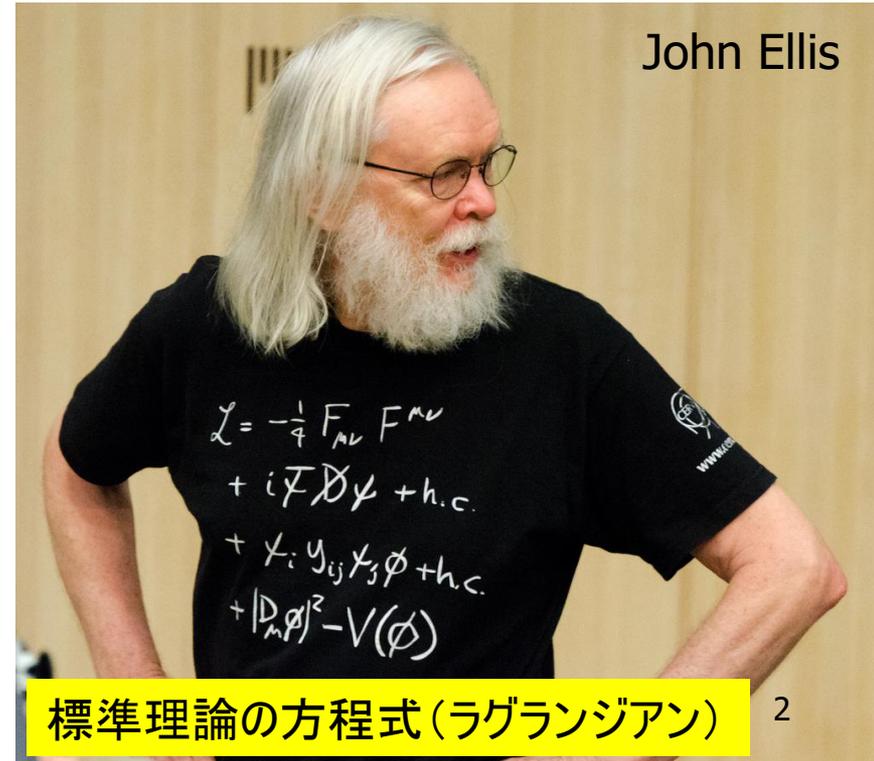
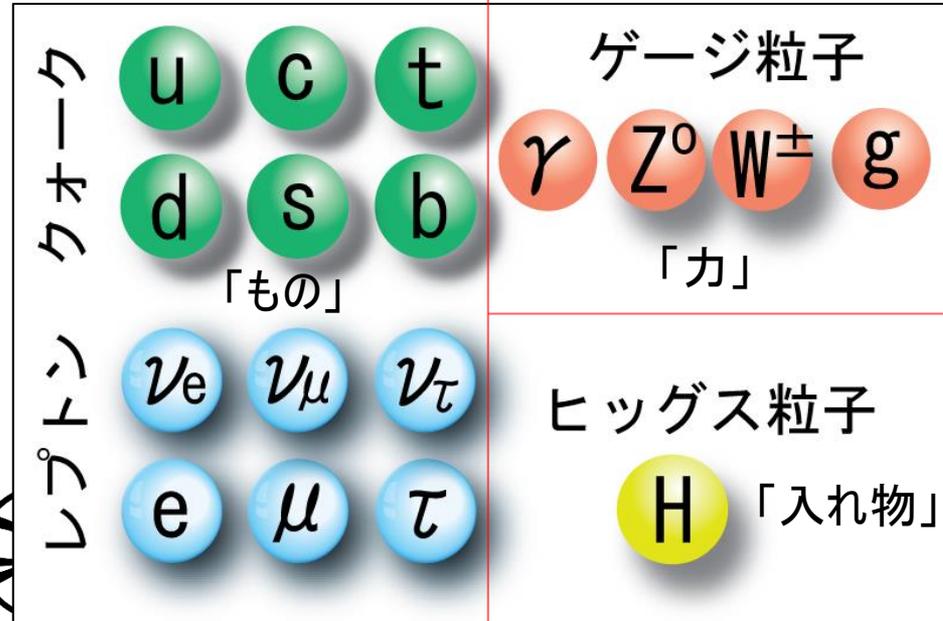
【共同通信】欧州核子衝突機関（CERN）のディレクター、ローフ・ヘイズ氏が、ヒッグス粒子の発見を報告する。ヘイズ氏は、ヒッグス粒子の発見が、素粒子物理学の歴史に新たなページを刻むと述べた。ヒッグス粒子は、物質の質量を生み出す役割を果たすと考えられている。CERNの大型ハドロン衝突炉（LHC）で発見された。発見は、理論物理学者フランソワ・エンゲルトとピーター・ヒッグスの予測と一致している。発見は、素粒子物理学の標準理論を支持する重要な証拠となる。発見は、素粒子物理学の歴史に新たなページを刻むと述べた。ヒッグス粒子は、物質の質量を生み出す役割を果たすと考えられている。CERNの大型ハドロン衝突炉（LHC）で発見された。発見は、理論物理学者フランソワ・エンゲルトとピーター・ヒッグスの予測と一致している。発見は、素粒子物理学の標準理論を支持する重要な証拠となる。



2013 NOBEL PRIZE IN PHYSICS
François Englert
Peter W. Higgs



全部で17個の素粒子



標準理論の方程式(ラグランジアン)

History of the Universe

現在

宇宙の背景輻射

現在

137億年

50億年

9億年

38万年

3分

10⁻¹⁰ 秒

10⁻³⁶ sec

10⁻³⁴ 秒

10⁻⁴³ 秒

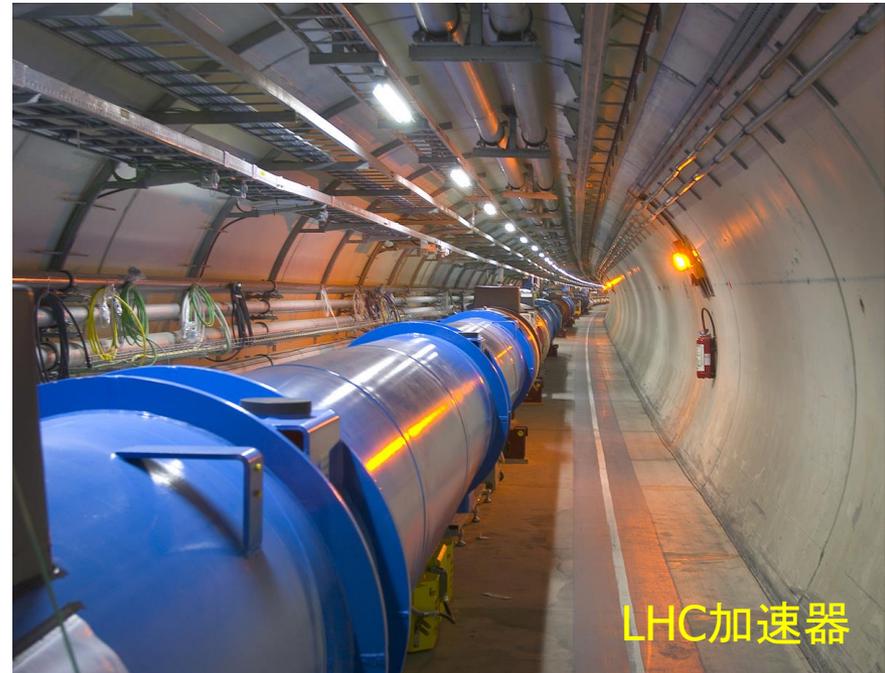
量子宇宙
10⁻³⁵ m

過去

素粒子実験 → 非常に「小さいもの」を測定する物理実験

実は「時間を巻き戻し」て初期宇宙を実現させている!!!

加速器: 粒子を超高速まで加速して
衝突させるマシン → 素粒子の生成
 $E=mc^2$



LHC加速器

世界最高エネルギーを誇るLHC加速器
13TeV → 開闢後10⁻¹²秒に迫る

標準理論では説明できない新現象・新粒子の探索 (例: 暗黒物質 → 超対称性粒子)

ジュラ山脈

Large Hadron Collider (LHC) ラージ ハドロン コライダー



アトラス

ATLAS

CERN: 欧州合同原子核研究機構

<https://www.cern.ch>

シーエムエス

CMS

LHCトンネル 27km, 地下100m

東京	山手線	34.5km
名古屋	名城線	26.4km
大阪	環状線	21.7km

ジュネーブ国際空港

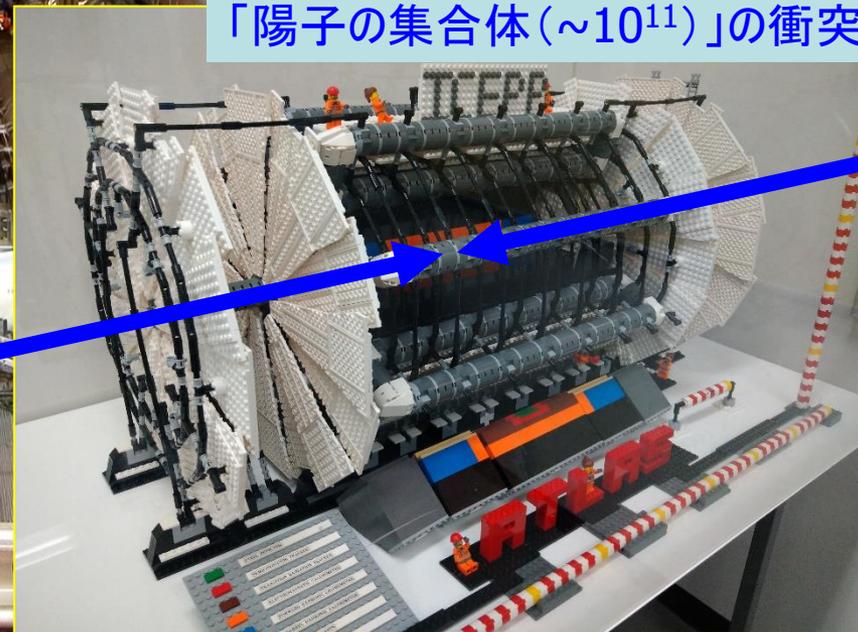


ATLAS

- ・高さ 22 m、長さ 44 m、重さ 7000 t
- ・世界最大の超伝導トロイド磁石
- ・センサー数 1.6億チャンネル

16000万画素のデジカメ

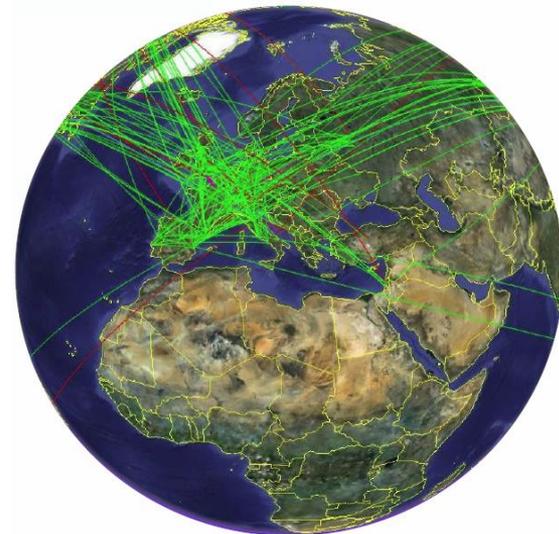
「陽子の集合体 ($\sim 10^{11}$)」と
「陽子の集合体 ($\sim 10^{11}$)」の衝突



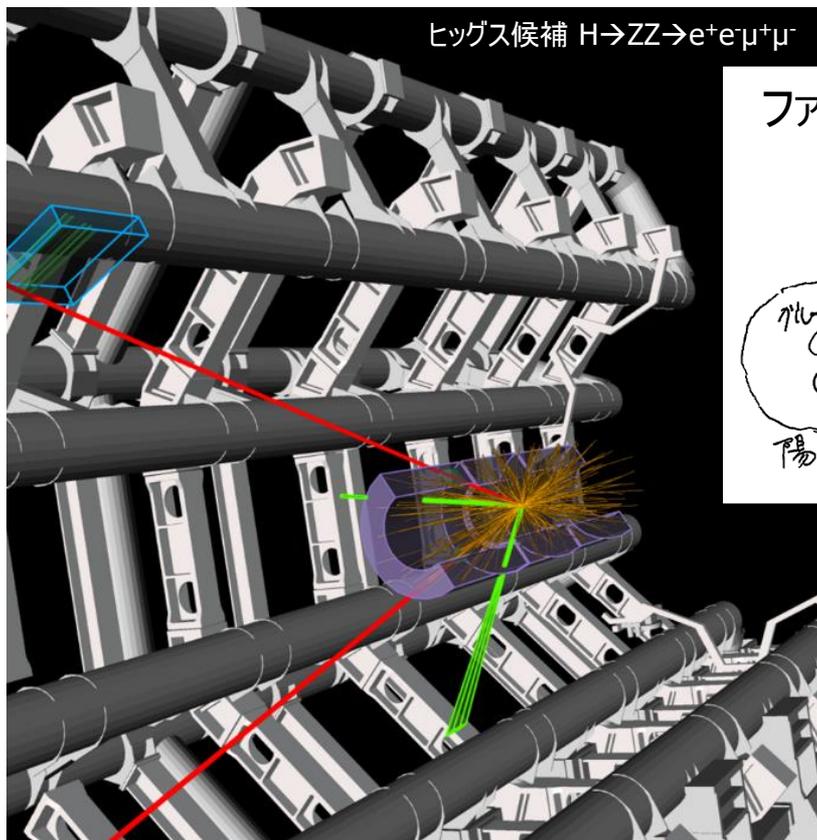
素粒子の研究とコンピュータ

WLCGモニター: データの流れ

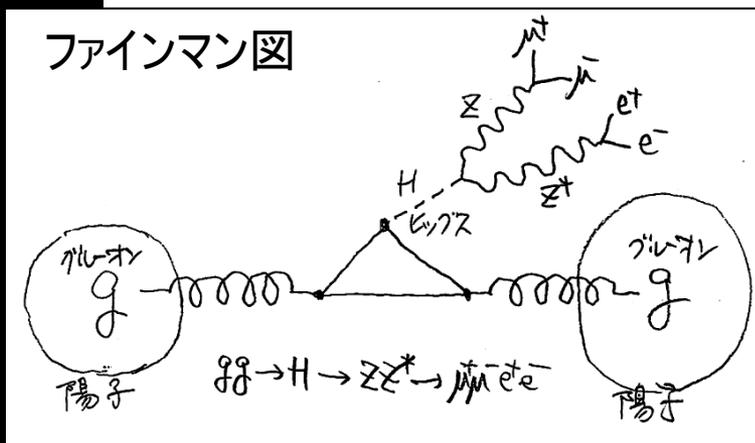
- 切っても切り離せない関係
 - 物理法則から導き出される「現象」を計算する。
 - 例) モンテカルロ・シミュレーション(MC)
 - ファインマン図
- WWWはCERNで誕生 → 研究者間の情報交換目的
- 実験データは「ビッグデータ」
 - ATLAS実験
 - 約400PB
 - 実験データ ... 200PB程度 (物理解析の一次データは約30PB~ 3×10^{10} 事象)
 - モンテカルロ・シミュレーションデータ(MCデータ) ... 200PB程度
- 「GRID」の実用化 → 世界中の高エネルギー物理学の計算機センターを一体化
 - Worldwide LHC Computing Grid (WLCG) <https://wlcg.web.cern.ch>
 - 日本は東京大学素粒子物理国際研究センターが計算機資源を提供
 - 学術情報ネットワークSINET5の国際回線のヘビーユーザー (欧州100Gbps、米国100Gbps)



モンテカルロ・シミュレーション データ



ヒッグス候補 $H \rightarrow ZZ \rightarrow e^+e\mu^+\mu^-$



ファインマン図

(実際にはもっと多くの図の計算を行う)



ファインマン

(1965年ノーベル物理学賞)

1事象の例: (たぶん)ヒッグス粒子が生成された事象

実験では、こういった事象が起こるかは確率.

→ 教師ありデータ

MC simulationでは粒子を「指定」して
特定の事象だけ大量に生成する.

→ 実験データと比較するため

年間 10^{10} のオーダー生成

1事象1分間必要なら 10^{10} 分~2万年かかる.

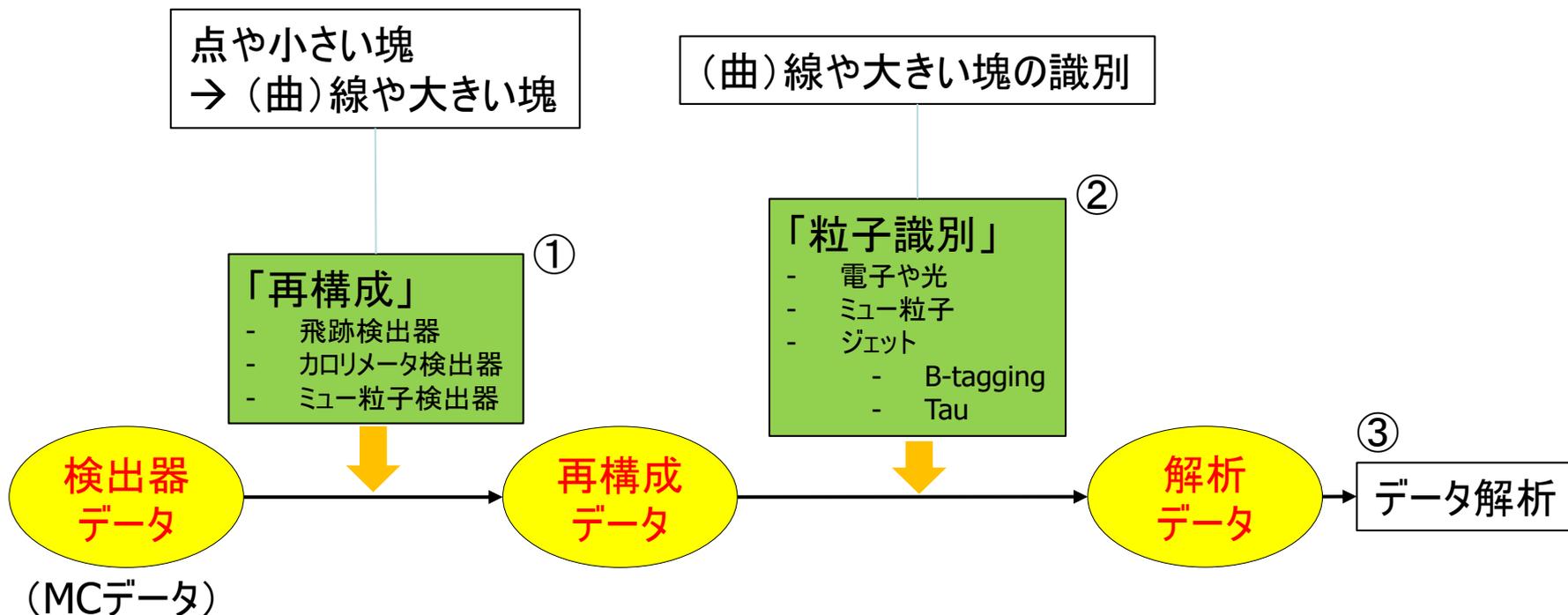
AI利用@素粒子実験

- 機械学習は以前から使われている。
 - 利用期間は長いが、実践ではまだまだ少ない(ブームと比較すると物足りない).
- 多くの場合 **Boosting+Decision Tree (BDT)** の手法
 - 我々の業界でのBDTブームの「きっかけ」は**2005年**のMiniBooNE実験の論文(らしい)
 - NIM A543 (2005) 577-584 (arXiv:physics/0408124)
 - “Boosted Decision Trees as an Alternative to Artificial Neural Networks for Particle Identification”

応用研究の目指すところ

- 4次元運動量(E, p_x, p_y, p_z)などの「**基本的な量**」の入力だけで従来の方法以上の性能を達成
- DLしか思いつかない**新しいアプローチ**の発見 → AIの理解
- **新しいDL手法**の開発

素粒子実験のデータの流れ



AIの利用場所

- ①再構成
- ②粒子識別
- ③データ解析

MiniBooNE実験は②
(②ではDLも活躍中)

以降①と③の話

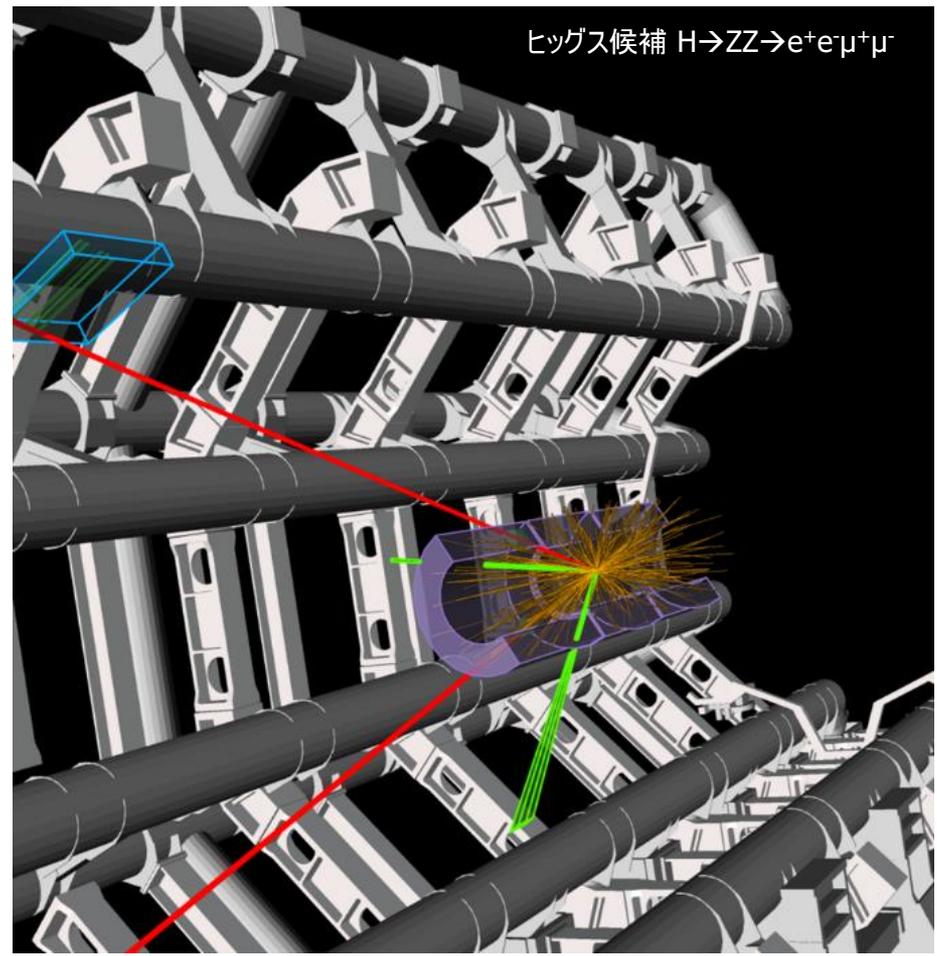
①

検出器データへのAIの活用

物理出力の向上を追求

→ **最新のDLやAI技術**を活用できる場所があるはず

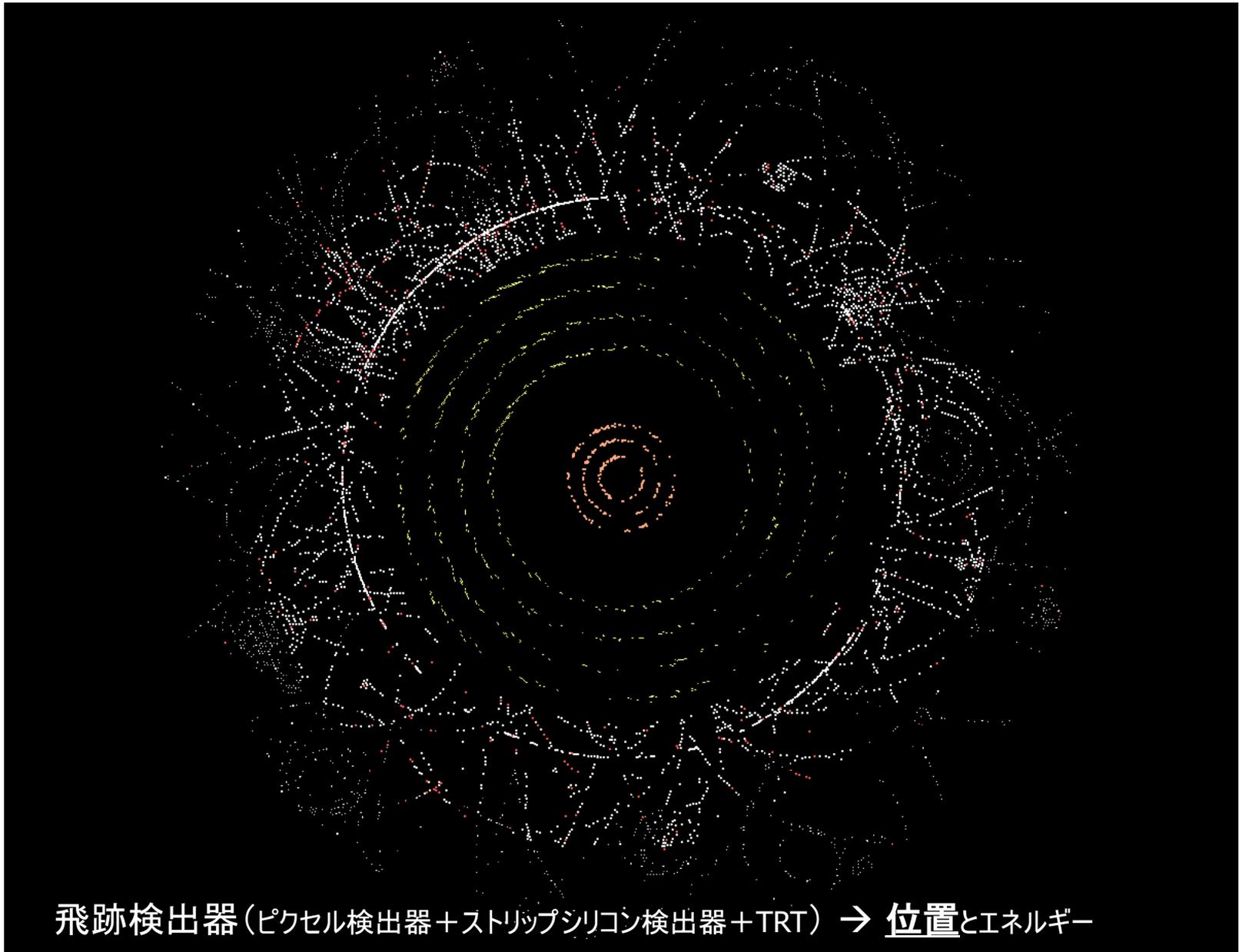
イベントディスプレイ =
「事象を綺麗に見せたもの」



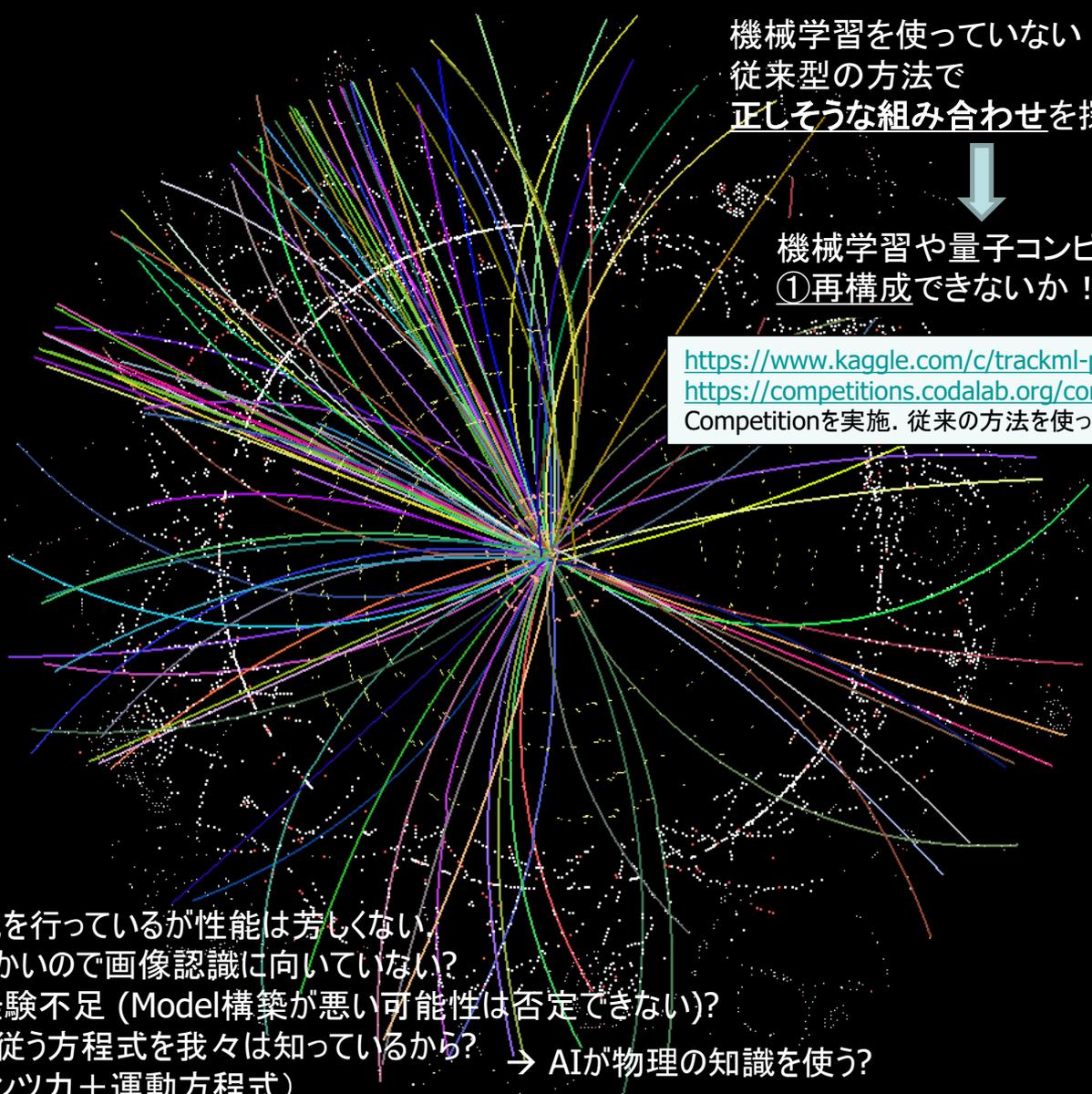
<https://twiki.cern.ch/twiki/bin/view/AtlasPublic/EventDisplayRun2Physics>

①

磁場がかかっています → 電荷をもった粒子は曲がる(ローレンツ力)



飛跡検出器 (ピクセル検出器 + ストリップシリコン検出器 + TRT) → 位置とエネルギー



機械学習を使っていない
従来型の方法で
正しそうな組み合わせを探した結果

↓

機械学習や量子コンピュータで
①再構成できないか！

<https://www.kaggle.com/c/trackml-particle-identification>
<https://competitions.codalab.org/competitions/20112>
Competitionを実施. 従来の方法を使ったアプローチが強い

我々も研究を行っているが性能は芳しくない

- 点が細かいので画像認識に向いていない?
or 経験不足 (Model構築が悪い可能性は否定できない)?
- 粒子が従う方程式を我々は知っているから? → AIが物理の知識を使う?
(ローレンツ力+運動方程式)

③

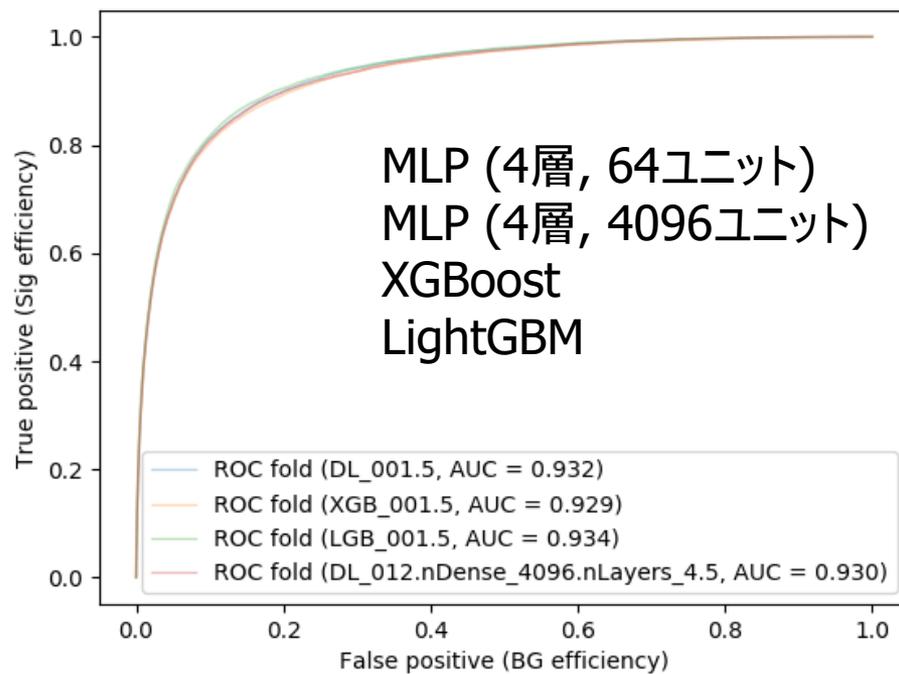
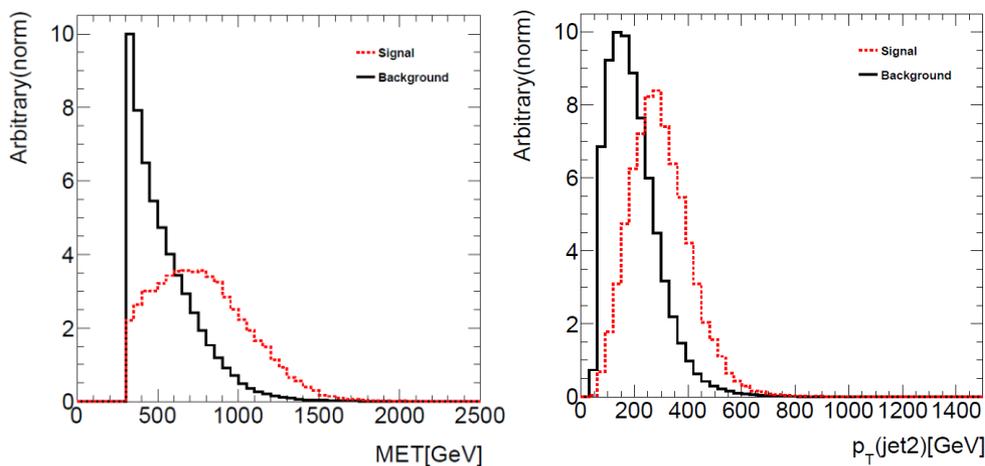
素粒子実験のデータ解析

③データ解析 BDTが主流

- Toolkit for Multivariate Data Analysis (TMVA) @ ROOT (2007)の整備
 - ROOT ... 高エネルギー物理学(HEP)業界の標準解析ツール, <https://root.cern.ch>
- せいぜい20変数を扱う程度なので、DLのご利益は今のところ見いだせていない(のかも).
 - 画像、映像、音声、言語などの応用例が適応しにくい。
 - MNIST → 28x28ピクセル = 784変数
 - ハイパーパラメータのTuning不足?

18変数の場合

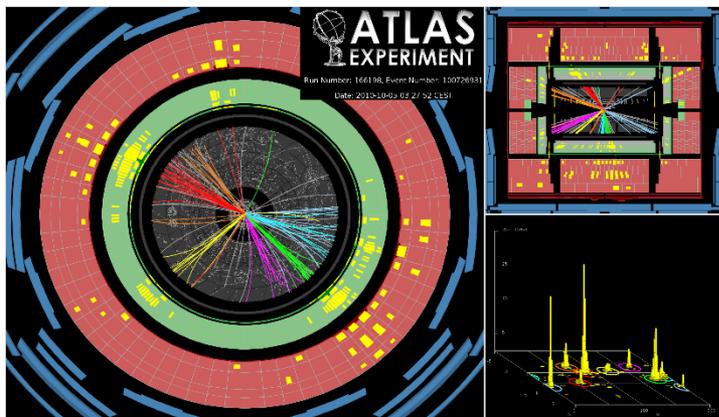
Signal=超対称性粒子の事象, Background=標準理論の事象



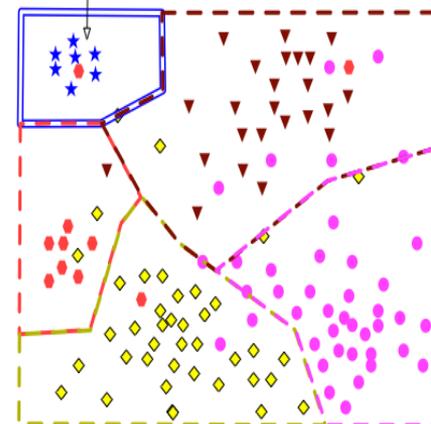
- DLでもBDTでも現時点では「物理法則」を使って人間が作った変数を入れた方がいい。
例: 質量 m (アインシュタインの有名な式: $E^2=(pc)^2+(mc^2)^2$)

③

自動「新粒子の探索」



新物理  異常検知



保持している事象

実験データ 約30,000,000,000

MCデータ 約100,000,000,000

データ解析では、ある程度「人間が決めた」条件で一部を選んでいる。
つまり、真剣に見ていないデータが大量にある。

→ 人間はすべてが標準理論だと信じている。

人間が見逃していた「小さな違い」からの発見を目指す。

「クラスタリング」や「オートエンコーダ」などを使って応用研究

- データの前処理(回転、反転など)、入力変数の選択などの研究

まとめ

- 素粒子物理学の大型実験

ビッグデータ → AIが活躍できる場・AIを成長させる場

- 様々な応用研究

- 再構成、粒子識別、データ解析

- DLが十分に活用できていない部分も多い.
 - 壁を越えれば、面白い知識が得られるはず.
 - 専門家の皆さんに相談したい

「物理学」と「AI」の繋がりから新しい道を!

<https://sites.google.com/view/icepp>

(すみません、現在Quantum Computingの方しか情報がないですが)