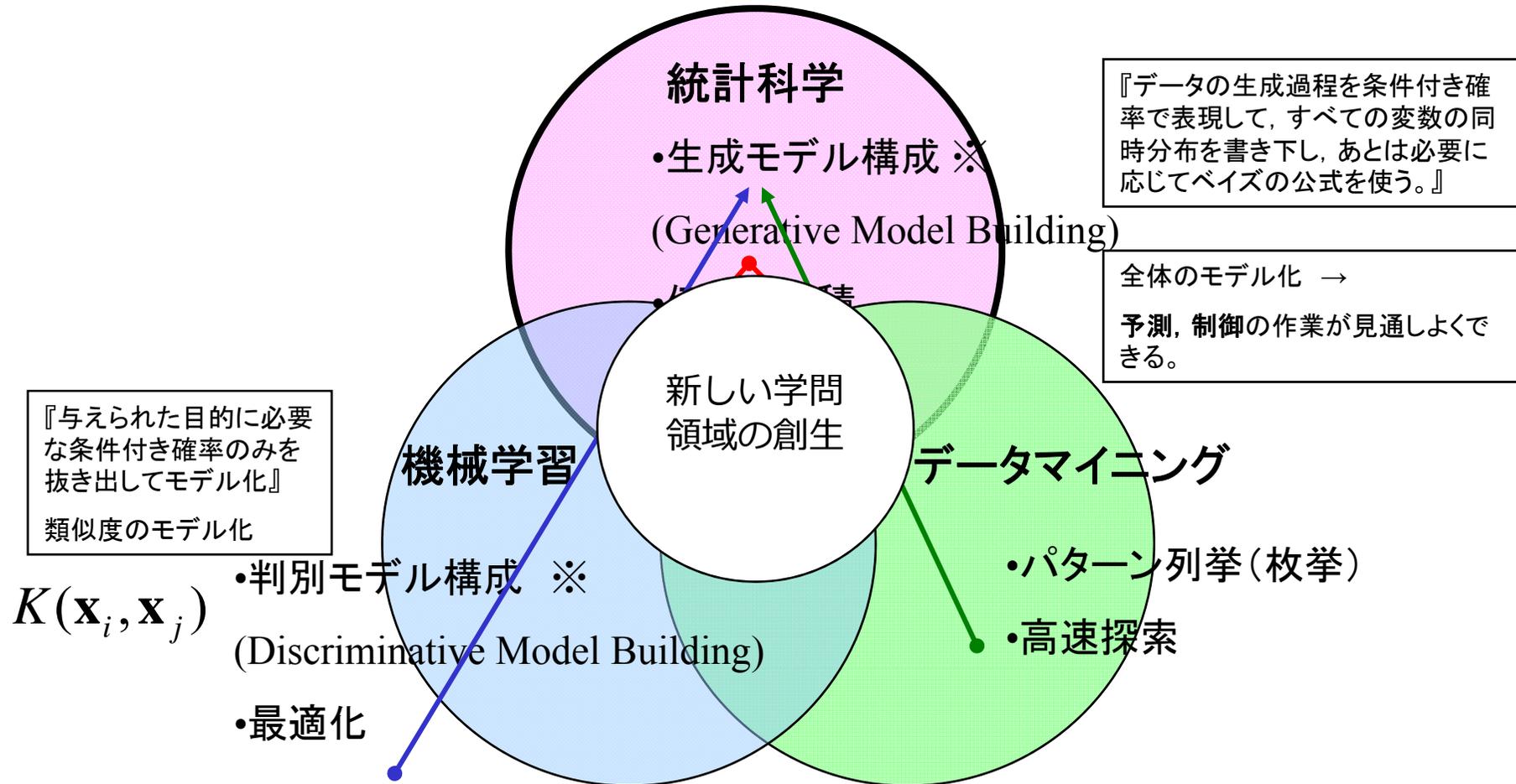


データ分析からうまれる、広がる 研究と交友の輪

樋口知之

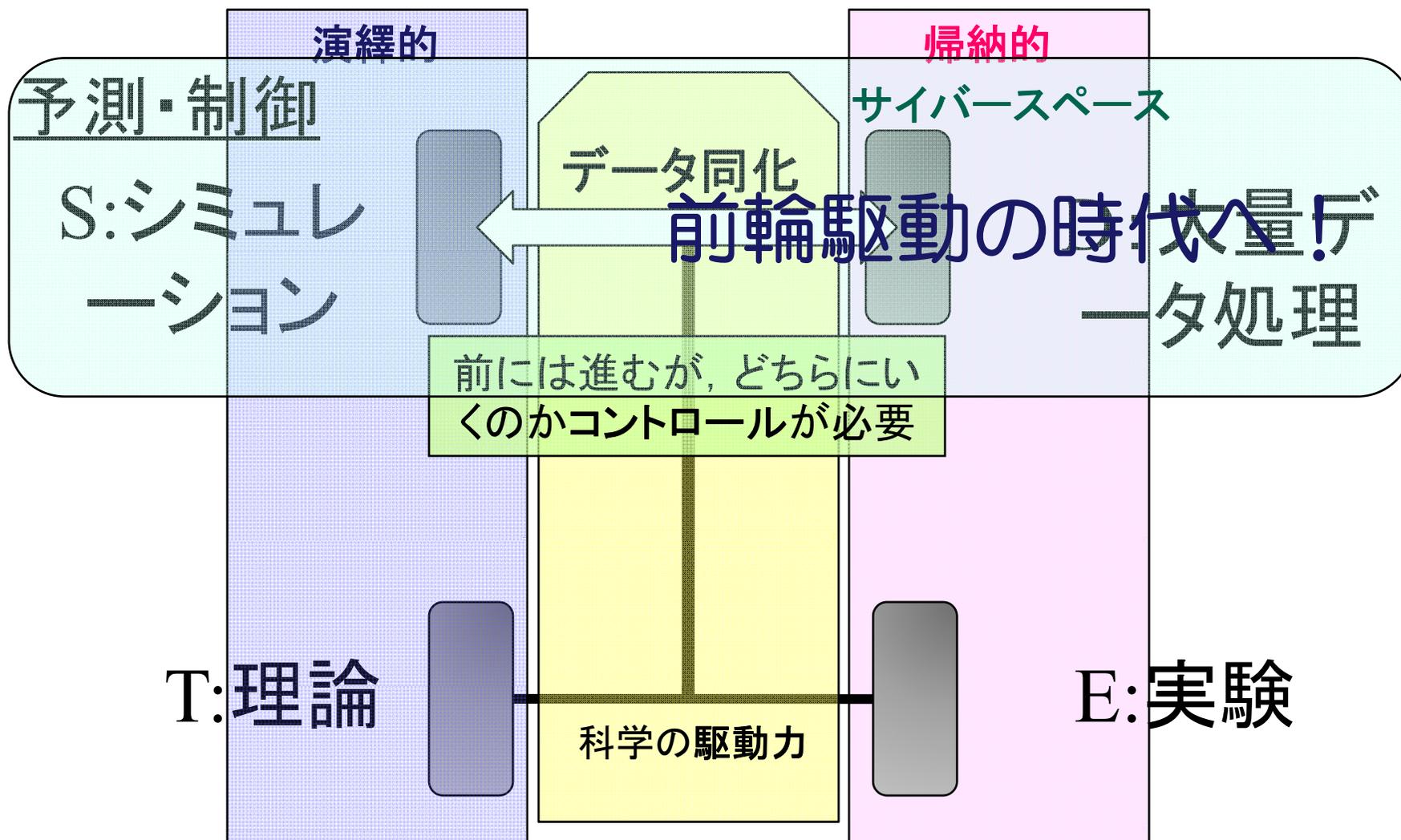
情報・システム研究機構 統計数理研究所 & JST CREST

超大量データ(情報)処理関連研究領域



※, 『 』: 伊庭による解説(信学技報告NC2006-55 (2006-10) 61—66)中の用語を利用

両者をつなぐのは禁手とされてきた



帰納的推論:「敵は本能寺にあり」

- 帰納的推論と機能のモデル化
 - 物理帝国主義観からの脱却
 - ニュートンパラダイムからのシフト
- 個人化技術 (*Personalization*)
 - 新しい np 問題 (NP 古くは Non-deterministic Polynomial time)

$n \ll p$ n : サンプル数、 p : 属性変数

機能のモデル化：“見よう見まね”を科学する

- “見よう見まね”のプロセスを加速する。
 - ― 体系化されていない研究分野において有効
- “見よう見まね”による完成の域がお手本を超える。
 - ― 贋作が“本物”を超える

日本人は伝統的に“贋作”を心から嫌う傾向が強いが、同じく漢字文化圏にある中国人で心ある人は、過去の文物に目を転じるとき、それがたとえ“贋作”であろうと判断できても、“本物”を超える出来栄であるならば、自分の心と目を満足させるため、“本物”以上にその“贋作”を尊重して手に入れると伝えられている。

“贋作”と“本物”の、どちらが本物か実は分からない

大量データ解析に必要な人材育成： モデラーの育成が急務

モデラーは

- 自分でモデルが作れる。
 - 自分で計算ができる。
 - 自分でモデル比較ができる。
- プログラミング言語を複数扱える。
 - 応用専門領域として主と副をもつ。
- チーム制の中で働ける。
 - 研究打ち合わせに熱心である。

■ サービス・イノベーションでも活躍できること間違いない