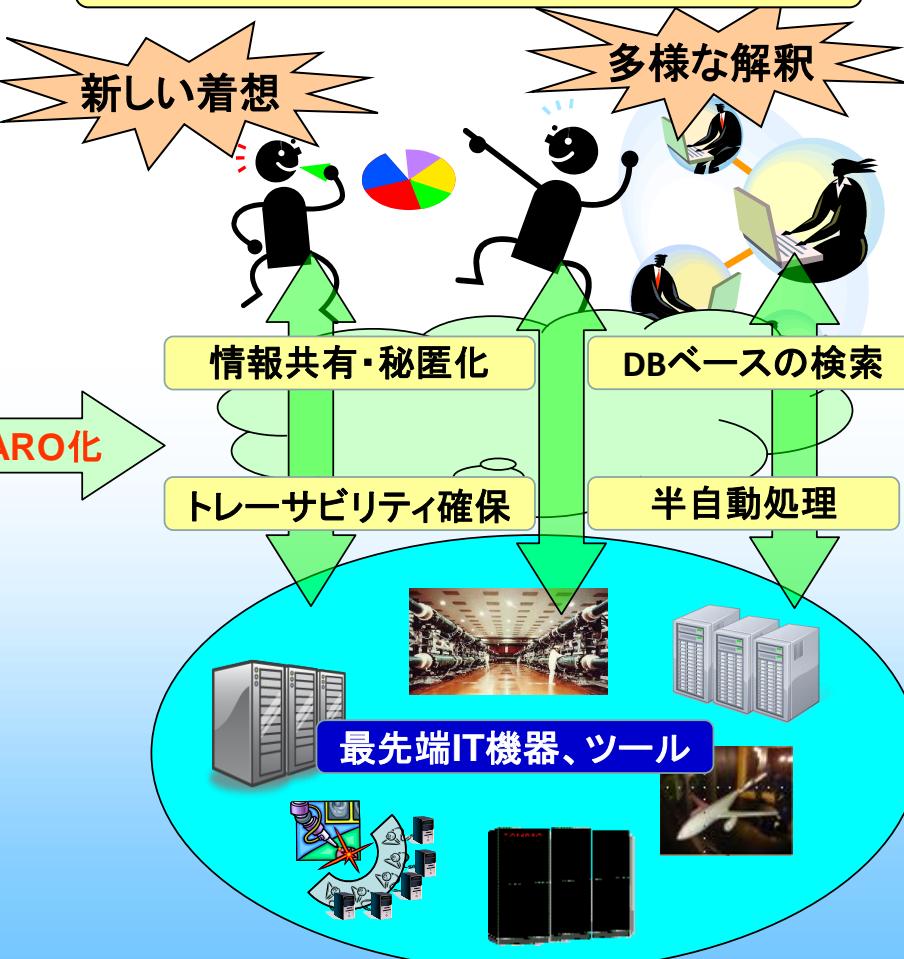


# 可視化、シミュレーションは、個人ツールの域を越えられないのか？

## 個人ツールとしてのシミュレーション



## 共有財としてのシミュレーション



# 目次

---

- 1) 会社紹介
- 2) なぜ、可視化/シミュレーションは、個人ツールの域を超えないのか？
- 3) 実験系では、共有、利活用が進んでいるのは、なぜか？
- 4) シミュレーションの共有、利活用は、本当に進まなくていいのか？
- 5) ASNAROで、可視化/シミュレーションを革新する！
- 6) まとめ

# 研究所で培ったR&D業務のシステム化ノウハウを結集し、 R&D統合プラットフォームを製品化しました。

## 社名の由来

紹介動画 : <http://kyoto-president.net/quatre>

**感性と知性を**

**情報技術で**

**革新します**

4つの i を科学する



・会社名	株式会社 キャトルアイ・サイエンス
・代表取締役	上島 豊
・本社	京都府京田辺市山手南2丁目1-3 ハチセンビル 3号館 3階
・設立	2006年1月
・資本金	22,500,000円

## 会社設立経緯

上島が研究員だった当時の課題(日本原子力研究開発機構でHPCによるシミュレーションを行う)

- 1) HPC性能向上によるデータ数・容量増大とそれらの処理・管理作業の増大
- 2) 属人的作業のミス発生や効率化の限界
- 3) 刻々と進展する結果及びコード、ツール共有の煩雑さ

これらの問題を解決する仕組みを研究員の時代に考案・開発

2000年 大規模データマネジメント研究会を立ち上げ

2003年 製薬会社や各種研究機関よりシステム相談を受け始める

2006年 キャトルアイ・サイエンスを設立

『研究開発を革新する』を原動力とし、R&D統合プラットフォームを製品化

販売開始から8年を経ており、R&Dプラットフォームとして、  
お客様から厚い信頼を得ております。

## 主要取引先(順不同、敬称略)

東京大学

大阪大学

北海道大学

東京工業大学

宇宙航空研究開発機構

日本原子力研究開発機構

防災科学技術研究所

高エネルギー加速器研究機構

情報・システム研究機構

情報通信研究機構

国立天文台

理化学研究所

計算科学振興財団

鉄道総合技術研究所

パナソニック株式会社

シャープ株式会社

株式会社神戸製鋼所

川崎重工業株式会社

新日鐵住金株式会社

中部地区自動車関連会社 等



Inspiration  
Intellect  
Information  
Innovation

4つの i を科学する

QUATRE-i SCIENCE

2)なぜ、可視化/シミュレーションは、  
個人ツールの域を超えないのか？

# シミュレーションは、「安価」、「やり直しがきく」、「古いコード・アプリの結果は不要」は、都市伝説！

「シミュレーションは、安価」というのは、計算機、ソフトウェア代、電気代また、関わっている人間全体の給与等コストおよび実験のデータ生成コストを客観的に把握していない発言。比較の為に、実験と同じデータを解析で作ることを考えると、全てのデータに関して、解析が「安価」というわけではない。例えば、実験で破壊検査的なものは、解析が比較的「安価」だが、非破壊検査的でパラメータでデータを振って傾向を見るものは、解析の方が「高価」である。

「シミュレーションは、やり直しがきく」というのは、データ管理がしっかりしていない以上、机上の空論。例えば、複数結果の3次元データの比較グラフで同じグラフが書けるかどうかを考えてみれば、多くの壁が横たわっていることがわかると思う。

「シミュレーションは、古いコード・アプリの結果は不要」というのは、データ管理がしっかりできておらず比較が困難であることを隠す為の言い訳。新しいコードの良さと悪さを理解するためには、常に比較するという観点が必要なはず。

「仕組み作りが困難＆面倒」  
「共有、利活用のメリットが無い」

「最低限の業務は、個人レベルで閉じて行える」が本当の理由！

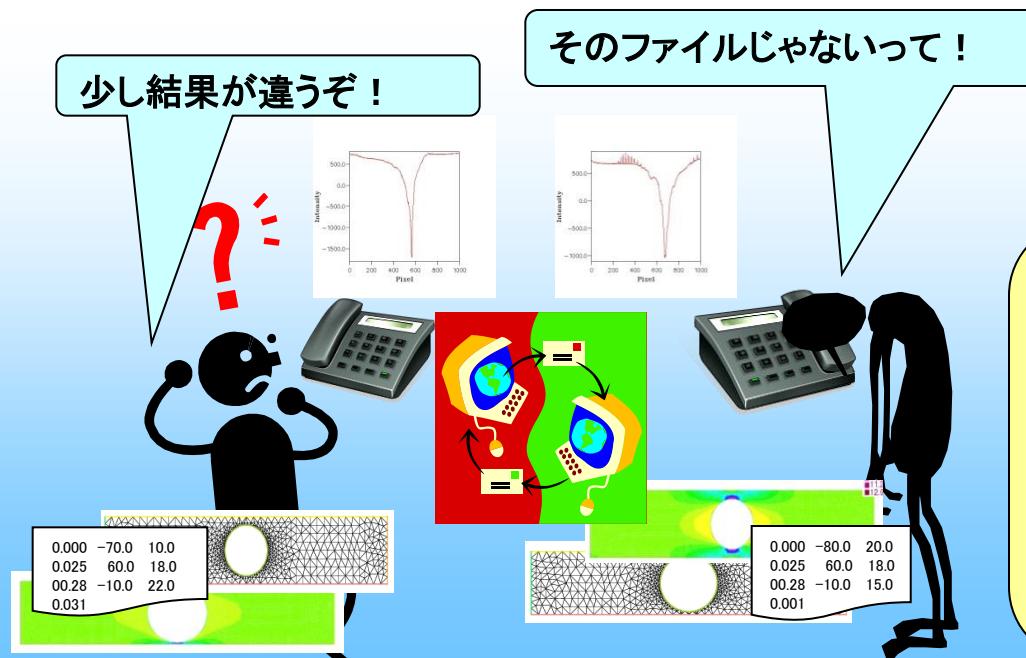
# データ・処理の紐付けが不完全！ 結果の元データや処理条件が、正確に分からぬ！

Who, What, Howを確実に！明確に！

解析結果から作ったグラフを再現するのに必要な情報がないと、第三者はその解析結果を正しく、把握できない。

(Overviewの場合はいいのだが、若干でも考察をしようとするところでは、グラフレベルで再現できる程度の情報が必要になるはず)

分からぬ部分を実際の解析実施者に聞くという行為は、「相手に迷惑がかかる」、「異なる立場の人々に自分の意図をうまく伝えられない」などで、精神的、時間的に高い壁がある。



# フリー(特に大学開発)のアプリ、ツールは、DLはできるが、 インストールが上手くいかない！

サポート、情報の  
貧弱さ



インストーラや設定UIの  
完成度

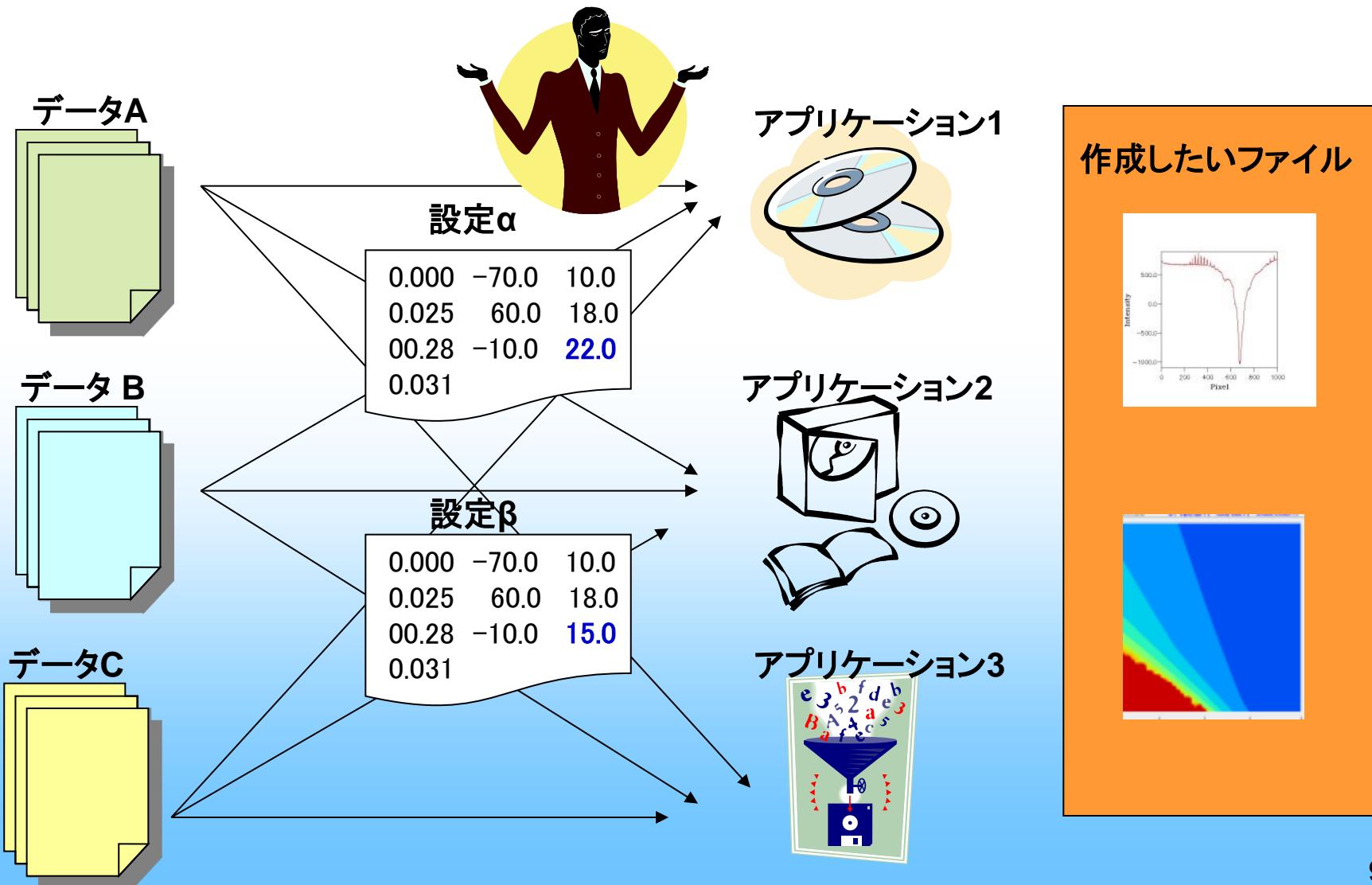


動作保証環境レンジの  
狭さ、あいまいさ



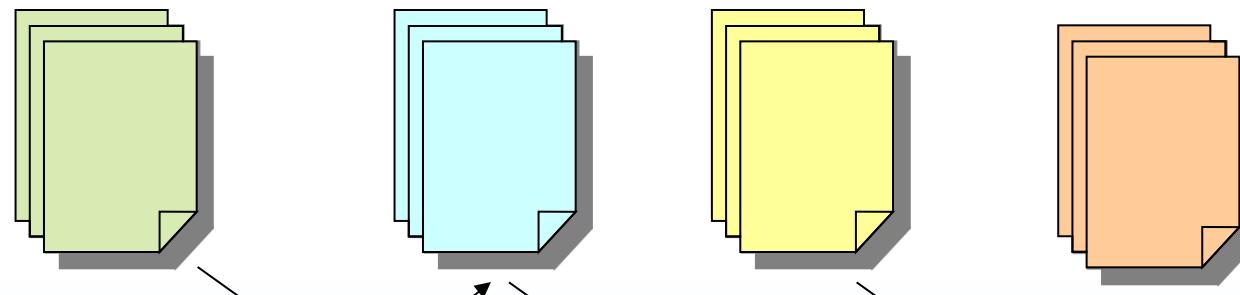
ソフトウェア、インストーラの  
バグ

# どのデータに対して、どのツールを、どのように使えば、良いか分からぬ！



# アプリ、ツールを動かせても、 正しく使っているのか不安で、それを確かめるのが困難！

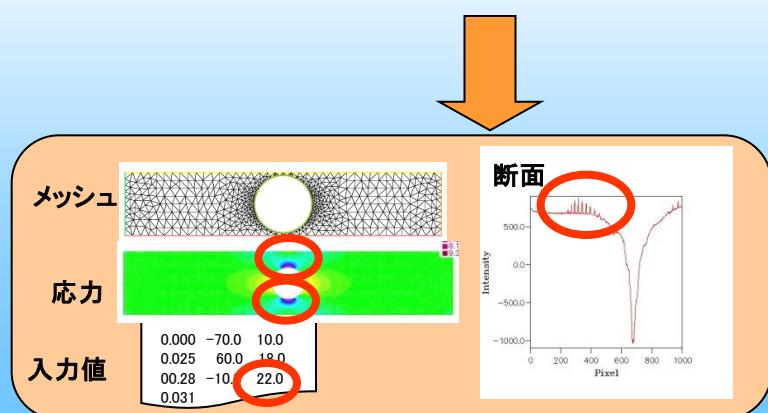
メッシュファイル 解析binaryファイル グラフ用ファイル グラフファイル



ソルバ

コンバータ

可視化ソフト



- ・ソルバが正しく動いていない？
- ・フォーマットコンバータがおかしい？
- ・グラフ化が間違っている？

3) 実験系では、共有、利活用が進んでいるは、  
なぜか？

# 大規模実験は、「最低限の業務でも個人レベルで閉じて行えない」ので「共有、利活用、仕組み作り」が必須！

「実験は、やり直しがきかない」というのは、精度問題である。実質的には、シミュレーションより、確実な再実験が可能なことが多い。同じ実験をすることより、他の実験をして、新しい情報を得る方が良いという意識から、再実験は避けるという判断になっている。

=>現状は、解析よりやり直しは確実！

「実験は、古い実験装置の結果は不要」なものとして、扱わない。理由は、新しい実験装置の良さと悪さを理解するためには、過去の結果も常に比較するという観点が必要だからである。

「実験は、高価」というのは、一部の実験に対してである。シミュレーションと同じデータを作ることを考えると、全てのデータに関して、「高価」というわけではない。例えば、実験で破壊検査的なものは、シミュレーションが比較的「安価」だが、非破壊検査的でパラメーターでデータを振って傾向を見るものは、実験が「安価」である。

「最低限の業務でも個人レベルで閉じて行えない」  
「共有、利活用が必須」

「実験 자체と比較すると仕組み作りは、困難というほどでない」

# 大規模実験では、計測器毎に評価したい量が決まり、 それ 자체の統計量に意味がある。

提供側



実験

設備の共有



データ、ツールの共有

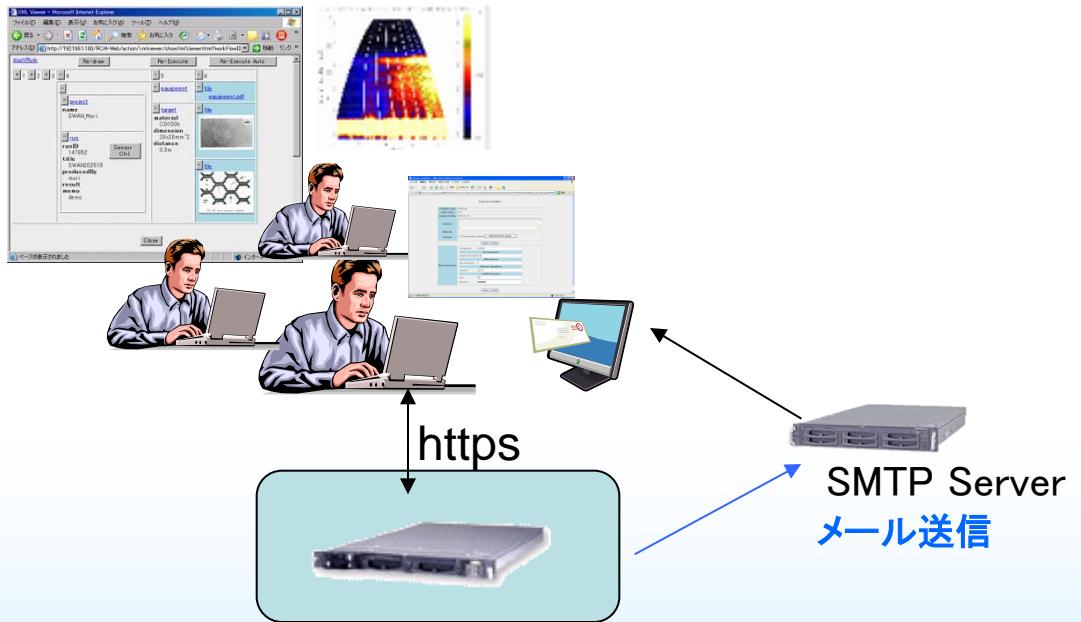
利用側

研究者  
分析ソフト

	評価したい情報	ターゲット	データ生成	データ処理
X線構造解析実験	計測器	ターゲット物質	実験実施	データ処理ツール
シミュレーション	ある程度は決めるが、 ポスト処理に任せると	計算格子	ソルバ実行	ポストアプリ

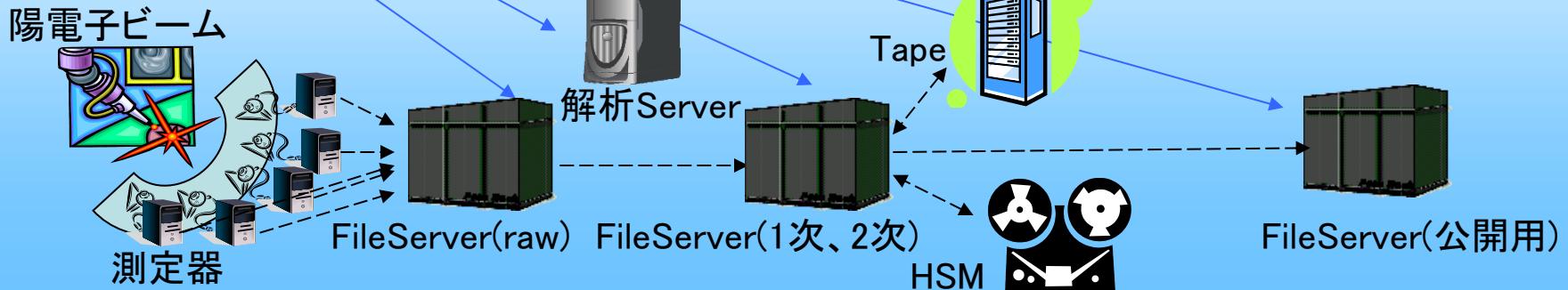
◎大規模実験でも最終的なデータ分析工程までは、共有、トレース管理はできていないが、生データは、すでにDB化されているので、システム化の道は、開かれている。

# 実験は初期データがDB化できているので、強力なDB化基盤さえあれば、複雑なデータ分析工程も共有化が可能！



## 効果

- ・自動データ処理と電子メールを連携した承認プロセスにより**正確かつ迅速なデータ処理**を実現
- ・大量のデータを**迅速に世界に発信**
- ・データの保存場所や統計情報の**メタ情報がDB化され、容易に検索**
- ・データ分析プロセスがトレースでき、考察時の**手間や不確実性を解消**



## 4) シミュレーションの共有、利活用は、 本当に 進まなくていいのか？

# すばる望遠鏡建設費は、地球シミュレータと同程度、運用費は1/3で、データ、ツールは共有化済！



参照:wikipedia:国立天文台ハワイ観測所すばる望遠鏡

建設・開発費	維持費	一日単価
約400億円	約20億円	約500万円



参照:wikipedia:地球シミュレータ

約400億円	約50億円	約1300万円
--------	-------	---------



約1100億円	約80億円	約2000万円
---------	-------	---------

参照:wikipedia:京(スーパーコンピュータ)

# 着手の容易さ、データの計測物理量、空間解像度は、 実験と比べて、シミュレーションにアドバンテージがある。

少し大きな実験は、実験装置の設定やモデル模型の作成などが必要で、全く新規の実験の着手ハードルが大きい。また、実験スケジュール調整や実験準備等があり、一人で勝手に実施できるものでは少ない。一方、シミュレーションは、少なくとも個人の判断で実施できることが多く、**着手ハードルは低い**。

実験は、計測器の設置の関係上、原理的に計測物理量や時空間解像度は、大きくできない。一方、**シミュレーションは、様々な物理量を高い時空間解像度で計測できる。**それゆえ、**高い自由度のデータ処理が可能である。**



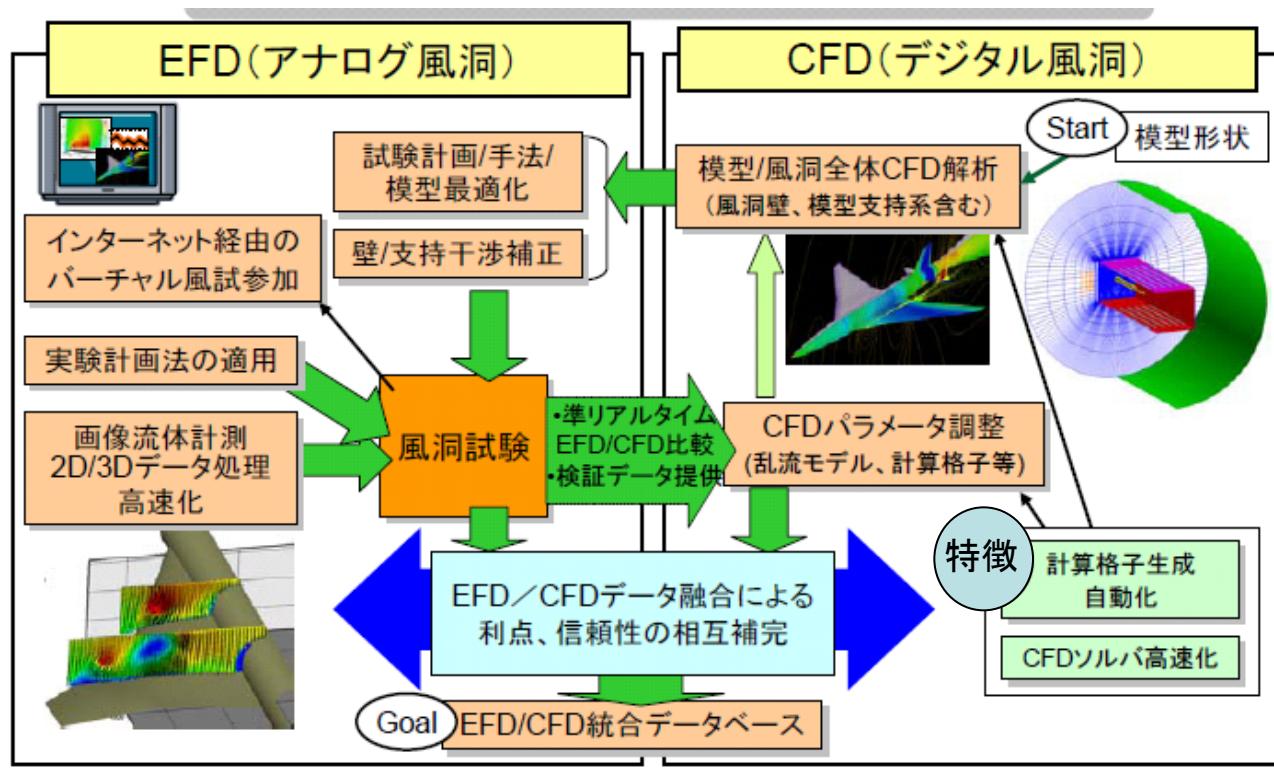
逆に言うと、シミュレーションは、単なる数値の統計では、必要な情報が得られなく、複雑なデータ分析が必須であり、それゆえ、**データ蓄積、プロセス記録が難しい！**

「着手ハードルの低さ」と「データ処理の高い自由度」は、  
シミュレーションの利点であるが、  
それが、データ蓄積を避け、再現性を低減させている原因でもある。

シミュレーションも蓄積していけば、実験と同様に 共有化  
メリットがあり、コードの強み・弱みも定量的に把握！

## JAXA DAHWINシステムでの適応事例

実験の弱点(壁・支持棒効果等)と計算の弱点(モデル適用範囲、精度等)を相互に補完！  
実験変更決定、コード改良指針、優先度も明確に！



### 効果

- 事前シミュレーションによる効率的な**実験条件の選定**
- 計算と実験の突き合わせ**による相互補完
- 遠隔地からの**風洞実験参加**
- 権限付与と承認フロー**による高度なセキュリティによる産学共同利用

精緻かつ多量の蓄積データに統計解析を適用することで、仮説の定量的検証や新たな発見が可能となる。

統計解析、学習システム等を適用する上で重要なこと

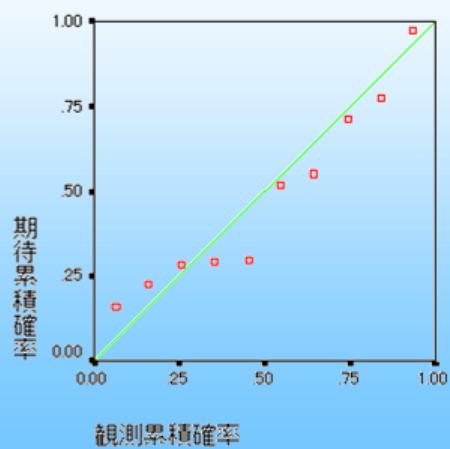
1)データの生成過程の完全な情報記録

=>データ母集団の品質問題

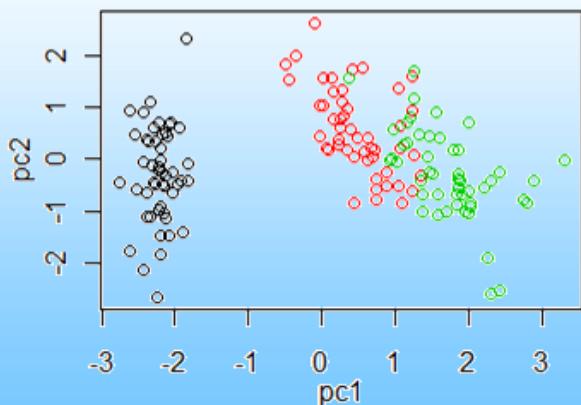
品質とは、情報に間違いが無いという意味と  
恣意性が働いていないという意味

2)データ母集団の量

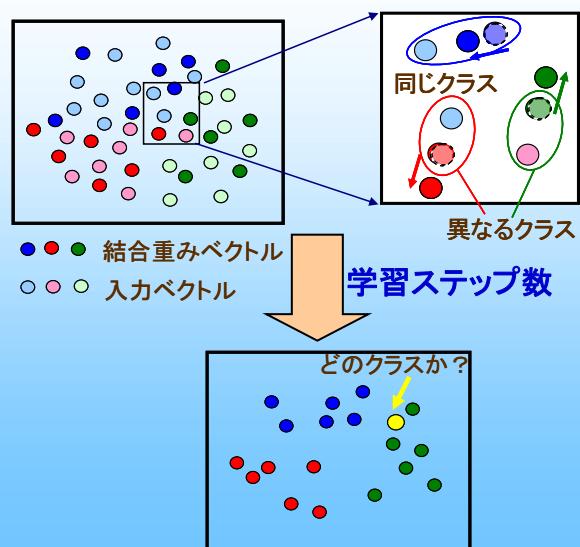
### 重回帰分析



### 主成分分析



### 学習ベクトル量化法



# 「データを蓄積することを諦めない！」気持ちさえあれば、シミュレーションも共有可能で、新たな知見が手に入る！

## a)研究開発的なシミュレーション

「研究開発的なシミュレーション」においては、結論を裏付けるためや発見プロセスのノウハウ継承のためには、詳細な解析プロセスの記録に基づく多様な視点からの検証が必要である。

## b)設計基準を満たしているかの確認シミュレーション

「設計基準を満たしているかの確認シミュレーション」においては、できる限り手動処理を排除した自動化により、評価の客観性、再現性を確保しつつ、解析ノウハウ、ルールの見直しが実業務データに基づき、逐次行われることが必要である。

「共有、利活用のメリットが無い」

=>前記のように多くのメリットがあるはず！

「最低限の業務は、個人レベルで閉じて行える」

=>もっと上を目指さないと実験再現のための解析止まり！

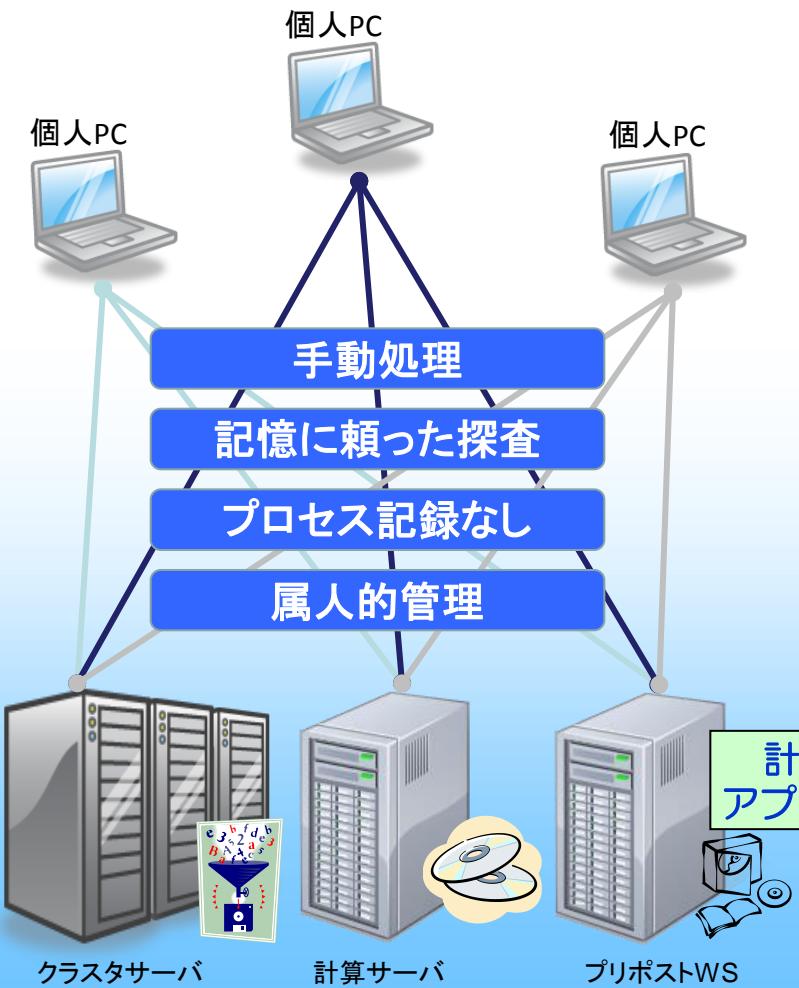
「仕組み作りが困難＆面倒」

=>弊社ソフトウェアASNAROで解決！

## 5) ASNAROで、可視化/シミュレーションを革新する！

# ASNAROは、開発フリーでシミュレーションコード、結果の共有、利活用を革新する！

## 従来の研究方式



## ASNAROによる研究方式



計算サーバ、  
アプリはそのまま

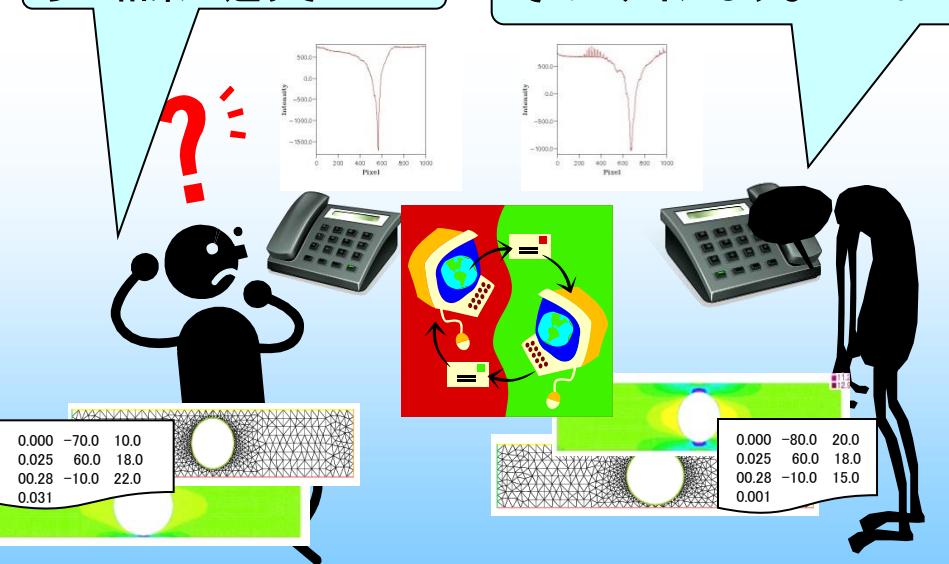
# ASNAROを使うことで、解析担当者の手を煩わせることなく、 解析を再現できるレベルの詳細情報が簡単に参照可能！

## 従来方式

電話、メール、対面で説明はしているのに…

少し結果が違うぞ！

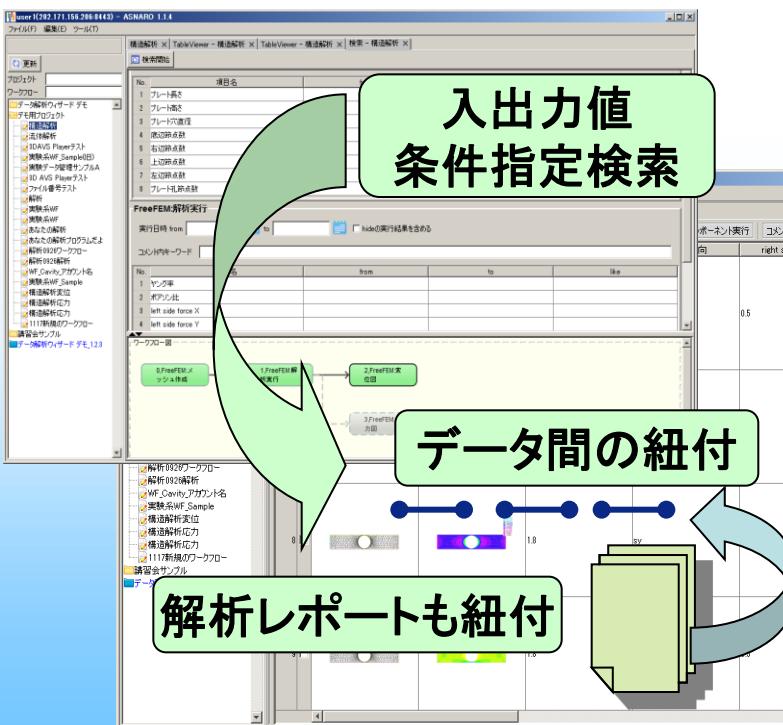
そのファイルじゃないって！



## ASNARO方式

解析再現に必要な全情報が検索一発！

入出力値  
条件指定検索



# 誰かがインストールしたものを使共用することで、 アプリ、ツールの整備ハードルは大きく低減！

従来方式

インストール自体が大変！

サポート、情報の  
貧弱さ

インストーラや設定UIの  
完成度



動作保証環境レンジの  
狭さ、あいまいさ

ソフトウェア、インストーラの  
バグ

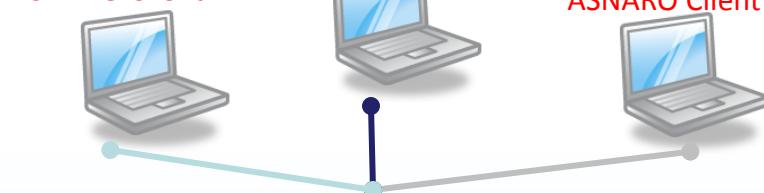
ASNARO方式

誰かがインストールしたらそれを共用利用！

個人PC  
ASNARO Client

個人PC  
ASNARO Client

個人PC  
ASNARO Client



ASNARO Server

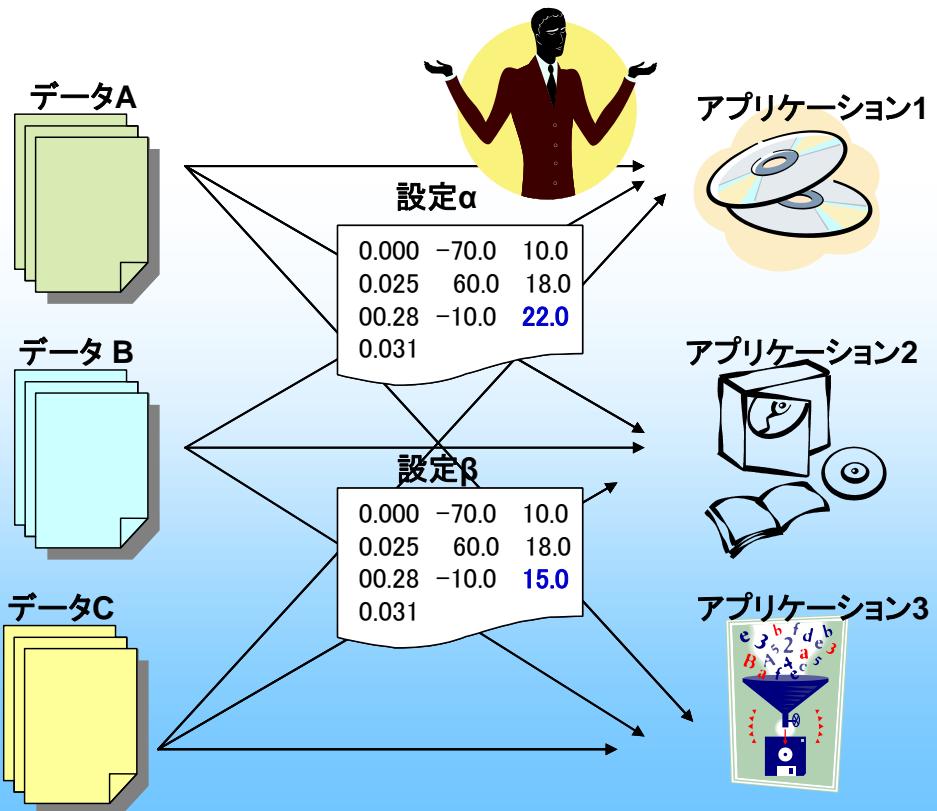
アプリ、結果の共用利用



# データとツールの紐付け機構により、どのデータには、どのツールを、どのように使えば良いかが一目瞭然！

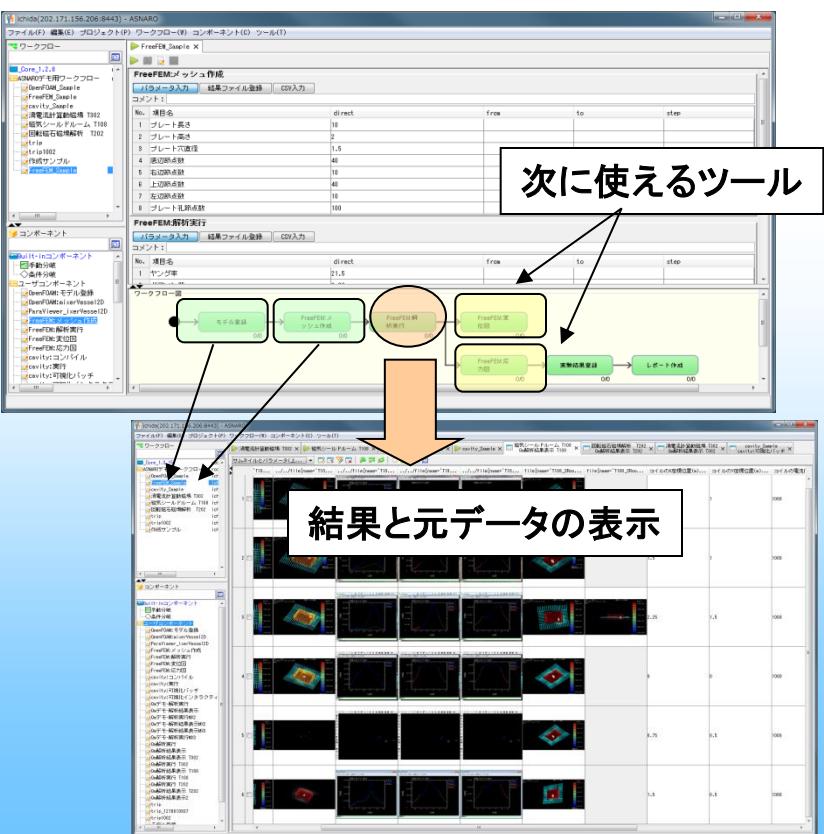
従来方式

どのツールが使えるのか分からぬ！



ASNARO方式

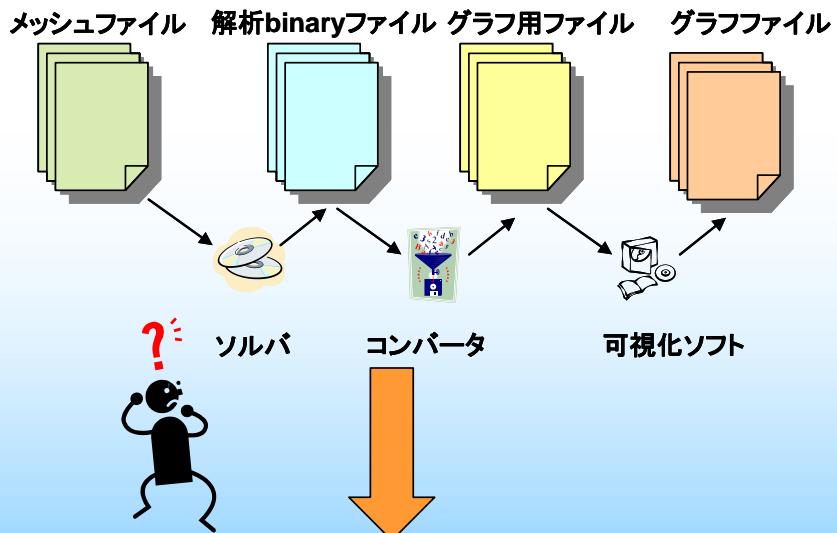
DBとWorkFlowでデータとツールが紐付け！



# データ生成のプロセス履歴が検索でき、表形式で並べられるので、類似結果から正しさを判断できる！

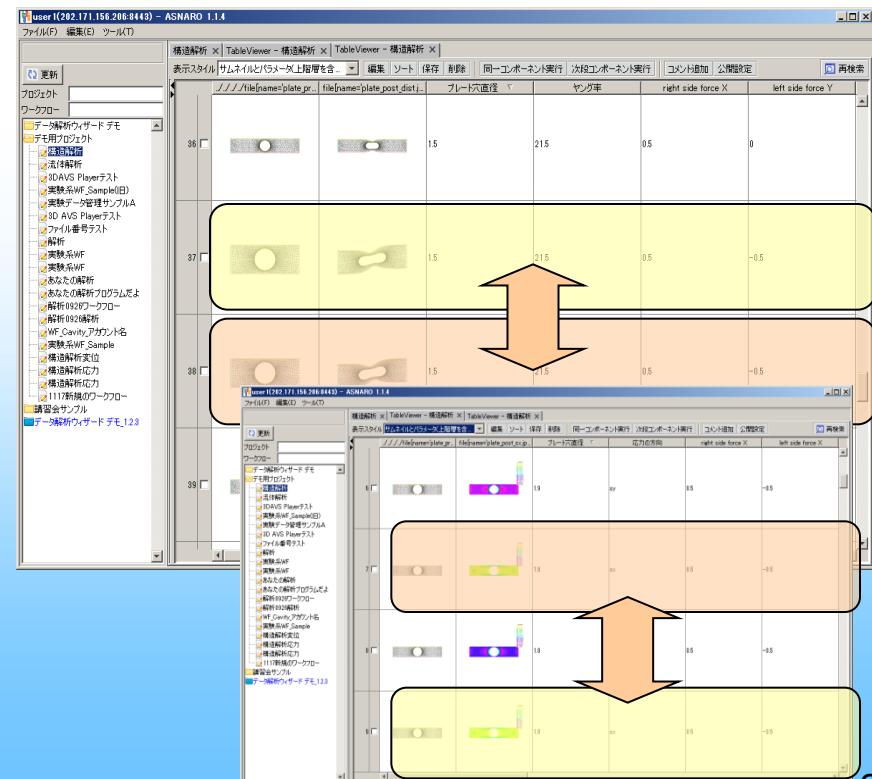
従来方式

結果が正しそうか判断ができない！



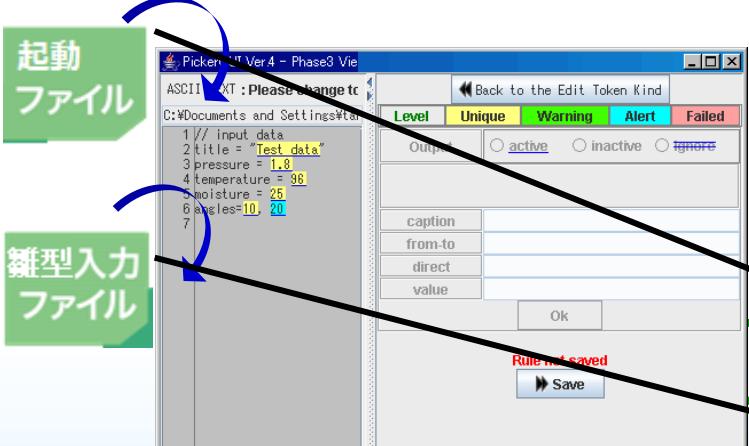
ASNARO方式

解析の過去類似結果から正しそうか判断ができる！

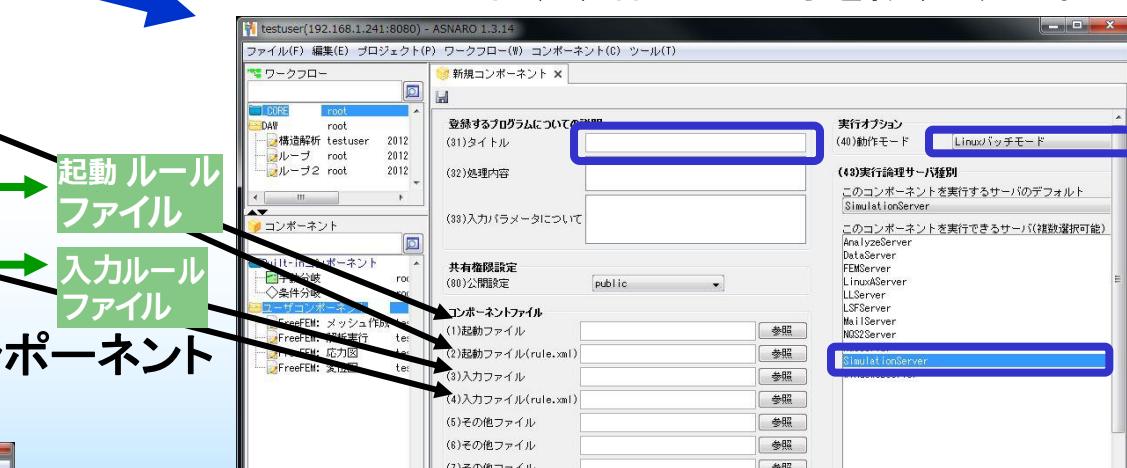


# Batchシェルと雛型入力ファイルが準備できたら、 3stepで、簡単に出入力値のDB化ができる。

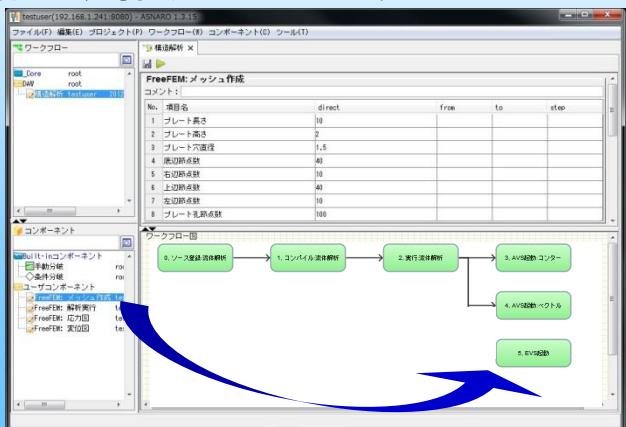
①Pickerツールで起動、雛型入力ファイル  
の変数化・文字列置換位置を指定する。



②コンポーネント登録画面で①ファイルを選択し、  
コンポーネント名、動作モード等を設定する。

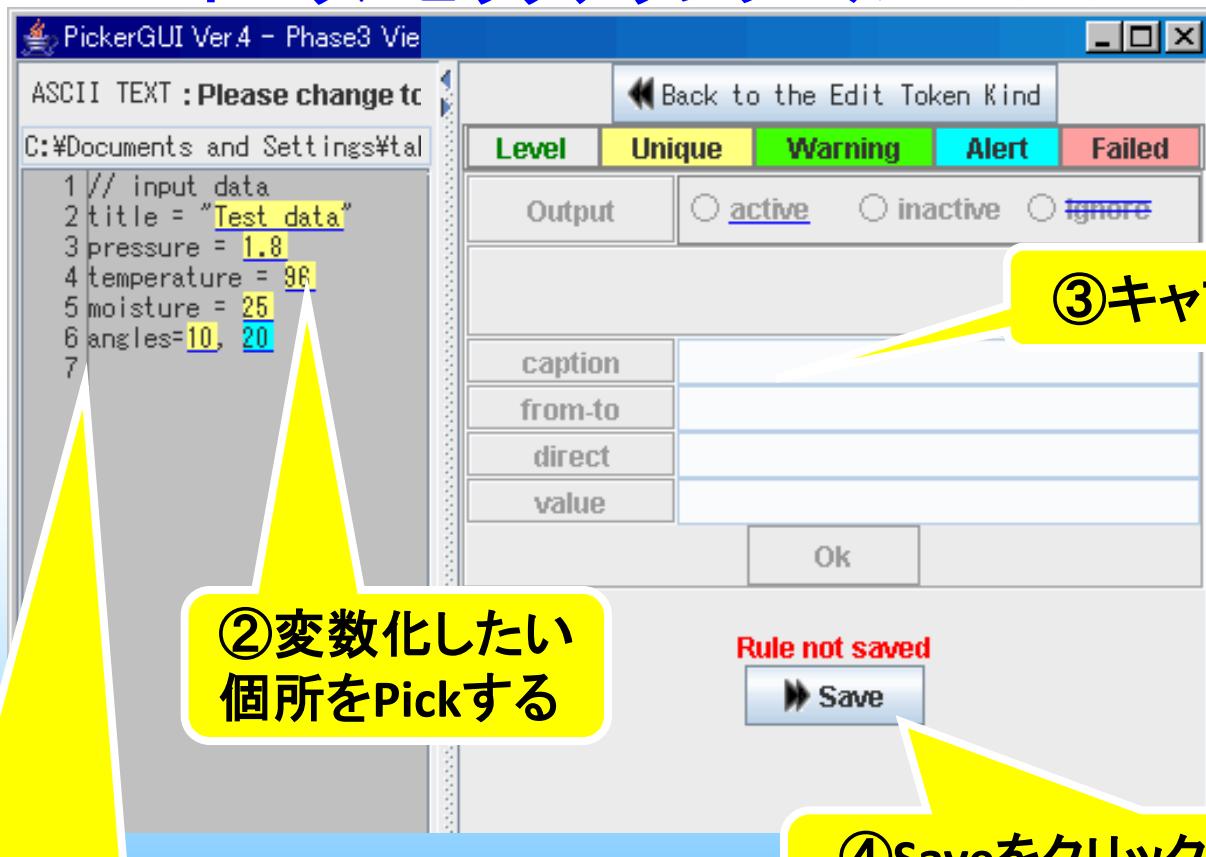


③ワークフロー編集画面で必要なコンポーネント  
を連結し、解析フローの完成



# 雛型入力ファイルと起動ファイルを準備し、 GUI化、DB化の為のルールファイルを作成する。

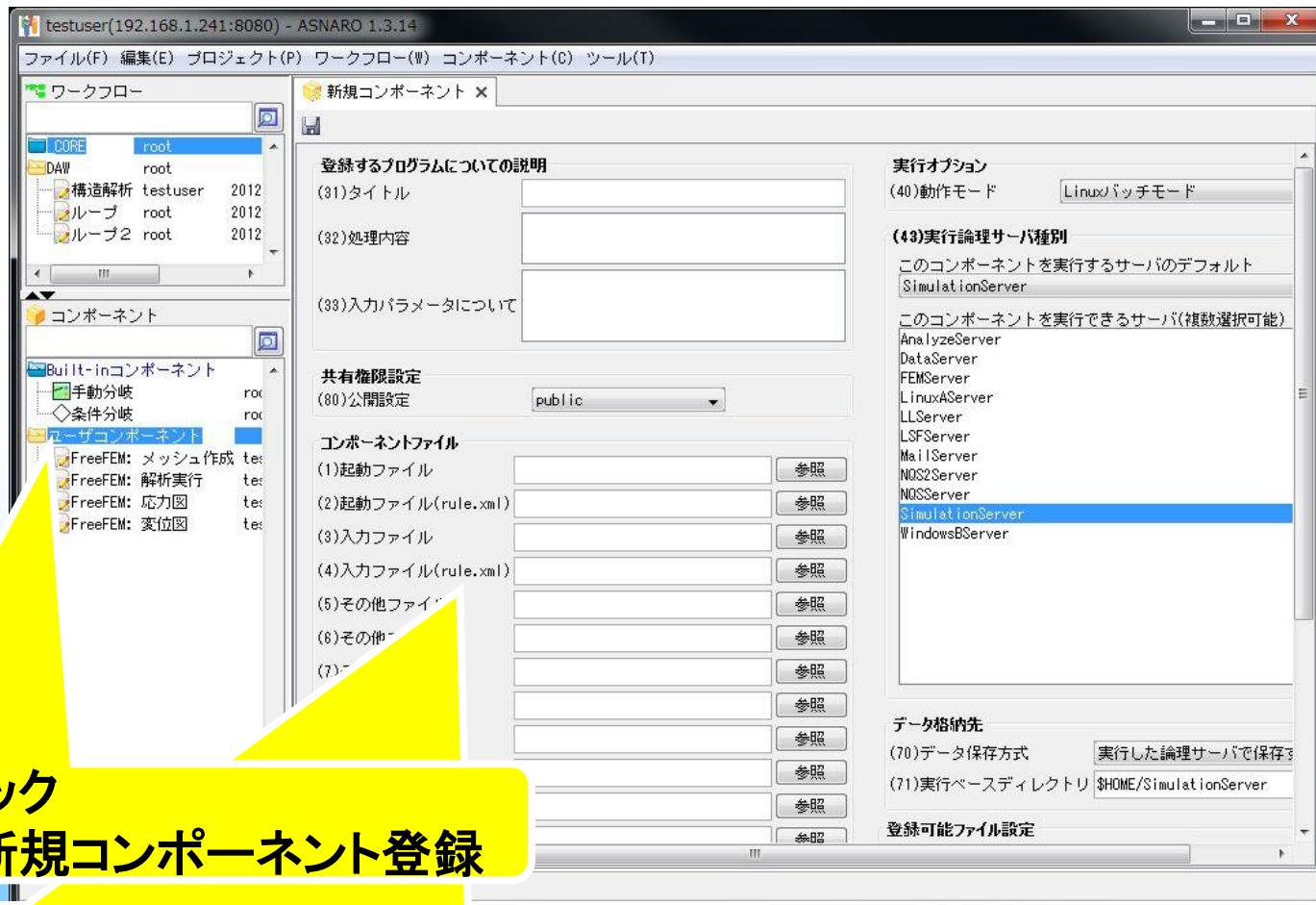
## トークンピックアップツール: Picker



①入力ファイルをLoadする

④Saveをクリックすると  
入力ルールファイルが作成できる

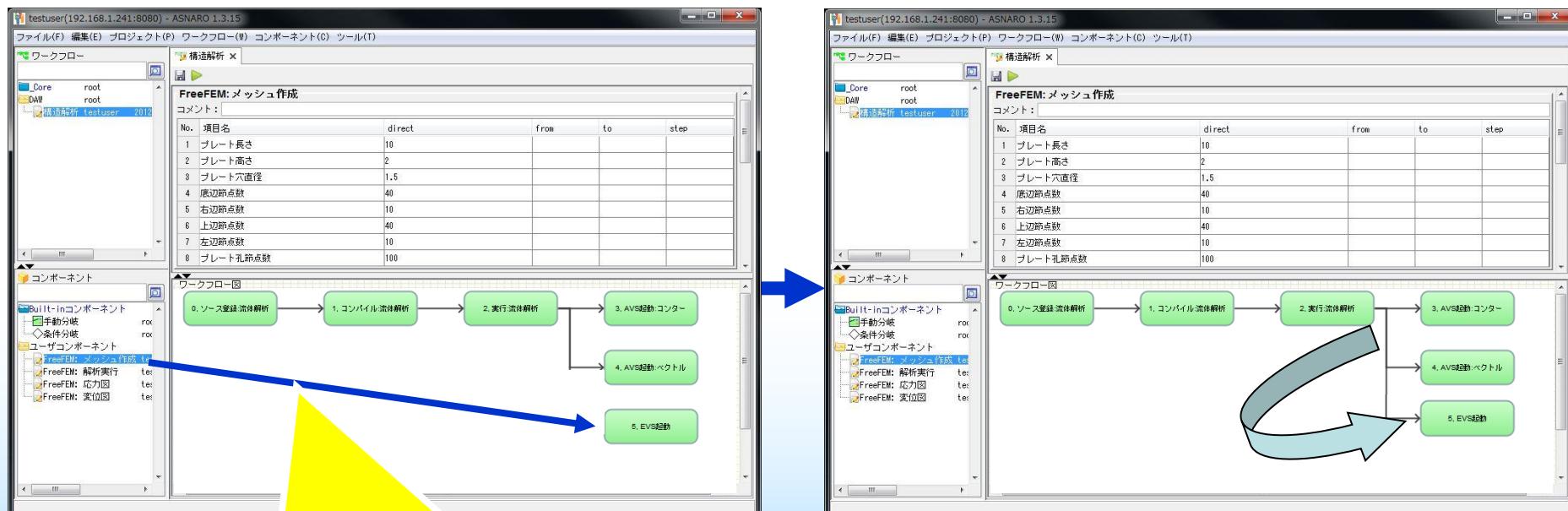
# 入力、起動、パターンマッチングファイルを選んで、コンポーネント名を入れるだけで登録は完了です！



①右クリック  
=>新規コンポーネント登録

②新規コンポーネント登録は、入力ファイルと起動ファイル、  
それらのパターンマッチングファイルをファイルセレクタで選択するだけ！

# 登録済コンポーネントをドラッグ&ドロップするだけで、新しい解析フローが簡単に作成できる。



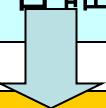
①登録済みのコンポーネントをドラッグ&ドロップするだけで、新しい解析フローが簡単に作成！類似解析フローの複製からの編集作成も可能！

## 6)まとめ

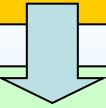
# 可視化/シミュレーションの個人ツール化の壁を破ることは可能！その先に新しい地平が広がっている！

シミュレーションデータ、ツールの共有、利活用により、新たな価値が創造される。

- データ、ツール利用の拡大で、多様な観点により、新たな発見が生まれる。
- ノウハウの共有・継承により、R&Dの安定的効率化が図られる。
- ツール、データ生成者外の第三者確認ができ、結論に信頼性が担保できる。



しかし、シミュレーションは、実験と異なり、その手軽さ・単独実施性および共有・利活用環境を整える困難さが共有、利活用の妨げになっていた。



シミュレーションの共有、利活用を進めるためには、実験と比べ、格段に簡単に共有・利用環境を整えられることが必要である。

## ASNAROは、それら問題を解決し、R&D環境を革新します。

- ツール、コード共有機構によるインストールハードルの低減
- データとツールの紐付け機構による処理手順、パターン選択ハードルの低減
- 完全なトレーサビリティ記録によるデータ生成プロセスを把握するハードルの低減

# 「ASNARO」は、R&Dプロセスを革新し、 技術者の創造性を支援します！

キヤトルアイ・サイエンスは、  
**感性と知性を  
情報技術で  
革新します！**

